Traceability of honey origin based on volatiles pattern processing by artificial neural networks☆

Tomas Cajka, Jana Hajslova *, Frantisek Pudil, Katerina Riddellova

Institute of Chemical Technology, Prague, Faculty of Food and Biochemical Technology, Department of Food Chemistry and Analysis, Technicka 5, 166 28 Prague 6, Czech Republic

ARTICLE INFO

ABSTRACT

Head-space solid-phase microextraction (HS-SPME)-based procedure, coupled to comprehensive two-dimensional gas chromatography–time-of-flight mass spectrometry (GC × GC–TOF-MS), was employed for fast characterisation of honey volatiles. In total, 374 samples were collected over two production seasons in Corsica ($n = 219$) and other European countries ($n = 155$) with the emphasis to confirm the authenticity of the honeys labelled as "Corsica" (protected denomination of origin region). For the chemometric analysis, artificial neural networks with multilayer perceptrons (ANN-MLP) were tested. The best prediction (94.5%) and classification (96.5%) abilities of the ANN-MLP model were obtained when the data from two honey harvests were aggregated in order to improve the model performance compared to separate year harvests.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

The European Union legislation (2001/110/EC) defines honey as "the natural sweet substance produced by *Apis mellifera* bees from the nectar of plants or from secretions of living parts of plants or excretions of plant-sucking insects on the living parts of plants, which the bees collect, transform by combining with specific substances of their own, deposit, dehydrate, store and leave in honeycombs to ripen and mature" [1]. Besides water, honey consists mainly of the monosaccharides (fructose and glucose) and many other substances such as organic acids, oligosaccharides, enzymes, vitamins, minerals, pigments, a wide range of aroma compounds, and solid particles derived from honey collection are present [2]. Honey is popular not only as a source of energy but also for its potentially health-promoting properties provided by prebiotic, antioxidant, antibacterial, and/or antimutagenic functionalities of certain constituents [3–6]. The price of honey is usually dictated by its botanical and/or geographical origin. While in the case of

botanical origin the most expensive are unifloral honeys, in the later case the higher price arises when honey is produced in a specific geographic location. Up to now, the EU has specified 18 protected denomination of origin (PDO) regions for honey (one Greek, one Italian, one Luxemburgian, one Polish, two French (including the island of Corsica), three Spanish, and nine Portuguese) [7]. Recently, an increased number of alerts concerning safety (presence of a variety of unauthorised or prohibited antimicrobial substances) and adulteration of honey have been posted [8–10]. In general, the adulteration techniques of honey are based on various principles: (i) water addition and extension with sugar and/or syrups; (ii) bee feeding with sugars and/or syrups or artificial honey; (iii) mislabelling as a results of mixing of honeys originating from different floral or geographical origin [2].

For the honey characterisation various parameters such as pollen analysis, moisture content, 5-(hydroxymethyl)furan-2-carbaldehyde concentration, sugar composition, proline content, invertase and diastase activity, and electrical conductivity are typically considered [2,11]. In addition to these traditional approaches, examination of the volatiles profile might be considered as a strategy enabling honey authentication since its composition (volatiles including) is known to vary widely with the floral origin and way of processing [12].

During recent years, solid-phase microextraction (SPME) in combination with the gas chromatographic–mass spectrometric (GC–MS) technique has been implemented as a method of

---

choice for the analysis of honey volatiles [13–18]. SPME, an inexpensive, solvent-free sampling technique enables convenient isolation of a wide range of low molecular weight analytes by their extraction from honey head-space (HS), or aqueous honey solution (direct immersion), and their concentration in the fibre coating. As regards GC–MS, due to the complexity of honey aroma, identification/quantification of its components can be hampered by co-elutions, which may occur when using conventional (one-dimensional) capillary GC set-up. Under these conditions, comprehensive two-dimensional gas chromatography (GC × GC) represents a challenging option permitting an efficient separation of all sample components [13]. In addition, a high-speed time-of-flight mass spectrometer (HS-TOF-MS) allows collection of the data at high acquisition rates (hundreds of spectra/s), required for proper reconstruction of very narrow peaks (50–500 ms) typically produced by GC × GC [19]. Also, due to a high TOF mass analyser efficiency, simultaneous acquisition of full mass spectra even at very low concentration of particular compounds is possible, what extends the possibility of identification of detected compounds on the basis of a library search [20,21]. It should be noted that rather higher cost of GC × GC–TOF-MS instrumentation compared to one-dimensional GC coupled to conventional MS detectors (*e.g.* quadrupole), is compensated by several advantages such as: (i) unbiased identification of sample components thanks to minimised co-elutions (hence better spectral quality); (ii) improved detection limits thanks to enhanced signal to noise ratio; (iii) faster GC separation (up to 4 times in the case of honey volatiles); (iv) possibility to shorten extraction time during SPME procedure, thus allowing an increase of the sample throughput significantly [13]. On the other hand, one should be aware of large volume of GC × GC–TOF-MS data, processing of which is rather demanding, unless smart chemometric analysis is employed. Typically, the principal component analysis (PCA) as a clustering method is applied for a preliminary inspection of the data structure, followed by the various classification methods such as linear discriminant analysis (LDA), discriminant partial least squares regression (DPLS), soft independent modelling of class analogy (SIMCA), or artificial neural networks (ANN) [22–26].

In this study, the application of HS-SPME–GC × GC–TOF-MS as an effective profiling tool, in combination with a chemometric approach employing ANN for data interpretation, is demonstrated with the emphasis to confirm the authenticity of the honeys labelled as "Corsica" (protected denomination of origin region) within the set of various European honey samples.

2. Experimental

2.1. Honey samples

In total, 374 honey samples were collected within the framework of the EU TRACE project [27]. In 2006 (first harvest), 111 Corsican, 18 non-Corsican–French, 15 Italian, 18 Austrian, 2 Irish, and 18 German honey samples were collected. During 2007 (second harvest), 108 Corsican, 28 non-Corsican–French, 15 Italian, 23 Austrian, and 18 German samples were collected. Before distribution, each honey sample was incubated at 40 °C overnight in an air oven, then manually stirred, and adjusted with distilled water to a content of solids of 70° Brix (harmonisation of measurement conditions for participants of the TRACE project). Prior to analysis, the honey sample (2 g) was placed into a 10-ml vial for SPME; after adding 2 ml of distilled water, the vial was sealed with a magnetic cap with PTFE/silicon septum and vortexed until complete homogenisation was achieved. (*Note*: Further dilution of the honey sample was important to enhance the transfer of volatiles from the honey solution to the head-space.)

2.2. Chemicals and materials

The SPME fibre 50/30 μm divinylbenzene/carboxen/polydimethylsiloxane (DVB/CAR/PDMS) used for sampling of honey volatiles was supplied by Supelco (Bellefonte, PA, USA). Prior to use, the fibre was conditioned following the manufacturer's recommendations.

The system used for GC × GC experiments comprised a DB-5 ms, 5% phenyl polysilphenylenesiloxane (J&W Scientific, Folsom, CA, USA) primary column; 30 m × 0.25 mm I.D., 0.25 μm film thickness, coupled via a column connector (Agilent, Palo Alto, CA, USA) to a SUPELCOWAX 10, polyethylene glycol (Supelco, Bellefonte, PA, USA) second column of dimension 1.25 m × 0.1 mm I.D., 0.1 μm film thickness. The upper temperature limits were 340 and 280 °C for DB-5ms and SUPELCOWAX 10, respectively.

A mixture of n-alkanes (C_8–C_{20}) dissolved in n-hexane employed for retention index determinations was supplied by Supelco (Bellefonte, PA, USA). The calculation was done for components eluting between n-octane and n-eicosane.

2.3. Instrumentation

A Pegasus 4D system consisting of an Agilent 6890N gas chromatograph equipped with a split/splitless injector (Agilent Technologies, Palo Alto, CA, USA), an MPS2 autosampler for automated SPME (Gerstel, Mülheim an der Ruhr, Germany), and a Pegasus III high-speed time-of-flight mass spectrometer (Leco Corp., St. Joseph, MI, USA) was used. Inside the GC oven a cryogenic modulator (N_2 jets–hot air jets technology) and a secondary oven (Leco Corp., St. Joseph, MI, USA) were mounted. Resistively heated air was used as a medium for hot jets, while cold jets were supplied by gaseous nitrogen cooled by liquid nitrogen.

The operating conditions of the optimised HS-SPME–GC × GC–TOF-MS method were as follows [13]:

(i) HS-SPME: incubation time: 5 min; incubation temperature: 40 °C; agitator speed: 500 rpm; extraction time: 20 min; desorption temperature: 250 °C; desorption time: 45 s (splitless). Once the splitless period finished, the injector was switched to the split mode (with a carrier gas flow of 50 ml/min) to remove any residues absorbed/adsorbed on the fiber. After 6 min exposure in the injector the fibre was automatically withdrawn and incubation and extraction of the next sample ensued.

(ii) GC × GC: primary oven temperature program: 45 °C (0.75 min), 10 °C/min to 200 °C, 30 °C/min to 245 °C (1.25 min); secondary oven temperature: +20 °C above the primary oven temperature; modulator offset: +35 °C above the primary oven temperature; modulation period: 3 s (hot pulse 0.6 s); carrier gas: helium (purity 99.9999%); column flow: 1.3 ml/min.

(iii) TOF-MS: electron ionisation mode (70 eV); ion source temperature: 220 °C; mass range: m/z 25–300; acquisition rate: 300 spectra/s; detector voltage: −1750 V (first harvest), −1500 V (second harvest).

ChromaTOF (LECO Corp.) software (v. 2.31) was used for instrument control, data acquisition, and data processing. Identification of compounds was based on a NIST 2005 mass spectra library search and was further confirmed by comparing linear retention indices available in the same library.

2.4. Chemometric analysis

Chemometric analysis included the principal component analysis and formation of an artificial neutral networks model employing the software package STATISTICA "Neural Networks" (v. 6, 2003, StatSoft, Inc., Tulsa, OK, USA) [28].

3. Results and discussion

In our earlier study [13], the HS-SPME–GC × GC–TOF-MS approach was optimised with emphasis on obtaining a high sample throughput and achieving chromatographic (high peak capacity of the GC × GC set-up) and analytical (spectral) resolution (deconvolution option for partially overlapped peaks) needed for good separation and identification of honey volatiles. To access various classes of compounds potentially present in honey head-space, a mixture of European honeys was used as a model sample for optimisation process. In this follow-up study, the optimised method has been used for examination of a large set of honeys of different geographical origin, thus, presumably, differing in profiles of the volatiles. The feasibility of using this assumption for a traceability purpose is presented in paragraphs below.

3.1. Characterisation of honey volatiles and selection of markers

Although only European honeys were involved in the examined set of samples, significant differences in both the extracted compounds and their concentrations were observed. Fig. 1 shows an example of GC × GC volatiles profile as obtained by analysis of one of Corsican honeys under earlier optimised HS-SPME–GC × GC–TOF-MS conditions. Using a 3 s modulation period (identified for the mixture of honeys as optimal), most of the volatiles were separated on the second (polar) column in a range of 0.5–1.7 s (second dimension retention time, ^2t_R), while only few of them were eluted with ^2t_R 1.7–3.0 s. Although, in this particular case, the separation space was not fully exploited, no further tuning of GC × GC system aimed at increasing the peak capacity was carried out since some honey samples contained polar volatiles with very high ^2t_R, thus, shorter modulation period would lead to their wrap-around and elution in the next modulation period.

For the chemometric analysis discussed below, several potential markers (volatiles) were selected (see Fig. 1) after careful examination of the GC × GC profiles of analysed honey samples. The selection criteria were large differences in intensities of the peaks among examined samples and known relationship with honey floral ori-
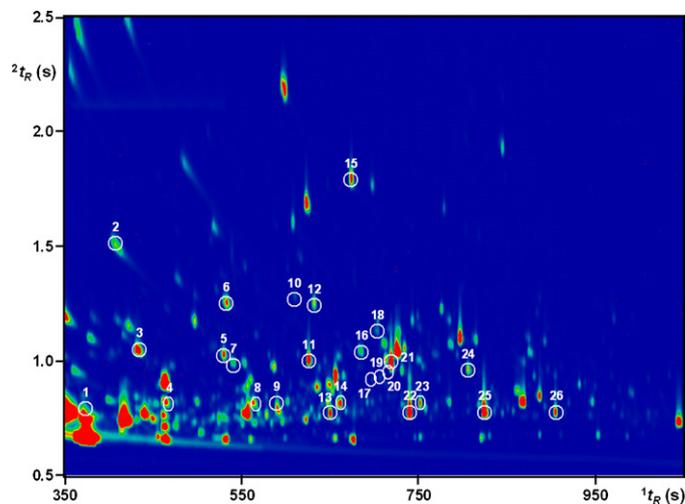

Fig. 1. HS-SPME–GC × GC–TOF-MS chromatogram of honey volatiles (Corsica sample) with marked markers. For the description of analytes see Table 1.

gin. These selected markers can be grouped as follows: (i) *aldehydes* (hexanal, heptanal, octanal, nonanal, decanal); (ii) *esters* (ethyl heptanoate, ethyl octanoate, ethyl nonanoate, ethyl decanoate); (iii) *sulphides* (methylsulfanyldisulfanylmethane); (iv) *alcohols* (hexan-1-ol, heptan-1-ol, octan-1-ol, nonan-1-ol, decan-1-ol); (v) *oxygenated aromatics* (benzaldehyde, 2-phenylacetaldehyde, 1-phenylethanone, 2-phenylethanol); (vi) *aromatic hydrocarbons* (1-methyl-4-propan-2-yl-benzene (*p*-cymene)); (vii) *ethers* (furan-2-carbaldehyde (furfural), lilac aldehyde I, lilac aldehyde II, lilac aldehyde III); (viii) *ketones* (3,5,5-trimethylcyclohex-2-en-1-one (isophorone), 2,6,6-trimethylcyclohex-2-ene-1,4-dione (4-oxoisophorone)). Interestingly, these selected volatiles are typically responsible for various flavour notes—aldehydes: green, sweet, citrus, floral, grape; esters: fruity, citrus, grape; alcohols: green; oxygenated aromatics: almond, floral, sweet, herbal, rose; aromatic hydrocarbons: citrus; ethers: woody, sweet, fruity, flowery; ketones: sweet, camphor-like, woody [29,30].

Table 1
Analytical data of selected honey volatiles (markers) used for chemometric analysis.

No.	Marker	^1t_R (s)	^2t_R (s)	RI	RSD (%)[a]
1	Hexanal	374	0.79	803	6.9
2	Furan-2-carbaldehyde (furfural)	410	1.49	836	3.9
3	Hexan-1-ol	434	1.04	870	5.1
4	Heptanal	467	0.81	903	8.2
5	Heptan-1-ol	530	1.02	972	4.8
6	Benzaldehyde	533	1.25	972	6.4
7	Methylsulfanyldisulfanylmethane (dimethyl trisulfide)	542	0.98	984	9.2
8	Octanal	566	0.81	1009	2.9
9	1-Methyl-4-propan-2-yl-benzene (*p*-cymene)	590	0.82	1034	4.7
10	2-Phenylacetaldehyde	611	1.27	1056	3.2
11	Octan-1-ol	626	1.00	1072	6.8
12	1-Phenylethanone	632	1.24	1078	2.1
13	Ethyl heptanoate	650	0.77	1068	12
14	Nonanal	662	0.81	1110	5.1
15	2-Phenylethanol	674	1.79	1123	5.7
16	3,5,5-Trimethylcyclohex-2-en-1-one (isophorone)	686	1.04	1134	4.8
17	Lilac aldehyde I	698	0.92	1148	9.1
18	2,6,6-Trimethylcyclohex-2-ene-1,4-dione (4-oxoisophorone)	704	1.13	1153	3.6
19	Lilac aldehyde II	704	0.93	1159	9.2
20	Lilac aldehyde III	716	0.96	1172	8.9
21	Nonan-1-ol	719	0.99	1173	6.8
22	Ethyl octanoate	740	0.77	1197	7.0
23	Decanal	752	0.81	1208	5.1
24	Decan-1-ol	806	0.96	1273	7.6
25	Ethyl nonanoate	824	0.77	1296	9.5
26	Ethyl decanoate	905	0.77	1396	10

[a] Relative standard deviation (RSD) of peak area, $n = 10$.

The list of the above compounds, their first (^1t_R) and second (^2t_R) dimension retention times, retention indices (RI) and repeatability of the measurement (expressed as relative standard deviation, RSD, %) is given in Table 1.

3.2. Chemometric analysis

Prior to the chemometric analysis employing PCA (unsupervised pattern recognition technique) and ANN (supervised pattern recognition technique) the raw data (374×26) presented in the form of absolute peak intensities (deconvoluted total ion current, DTIC) were pre-processed. In the first step, they were scaled using the range transformation, *i.e.* the lowest value of given variable was assigned to "0" and the highest one to "1", whereas the remaining entries were numbers between these values within the interval (0,1). This procedure transformed all data to a uniform range of variability. In the next step, logarithmic transformation of the data was carried out. A general advantage of this transformation was reduction of differences in variation of experimental data [26,31]. In our study, the differences of absolute intensities of particular volatiles ranged between 2 and 5 orders of magnitude. Under these conditions, transformation is important to support the contribution of even small variables (small GC peaks in particular case) to the classification of samples.

3.2.1. Principal component analysis

Principal component analysis (PCA) represents one of the most frequently used chemometric tools mainly due to its very attractive features. PCA allows relatively easy projecting of data from a higher to a lower dimensional space and then reconstructing them without any preliminary assumptions about their distribution [32]. In the preliminary data analysis, PCA was performed to investigate any possible clustering of samples on the basis of geographical origin. As Fig. 2 shows, using PCA, the honeys were divided into two groups described as "Corsica" and "non-Corsica". The first principal component (PC1) accounted for 26.0% and 28.8% variance for year 2006 and 2007, respectively, while the second principal component (PC2) contributed for 14.7% and 18.4% for these years. For merged data generated in years 2006 and 2007, the PC1 and PC2 were 27.2% and 16.3%, respectively. Considering the eigenvalues >1, seven, the most important, PCs contributed to 77% of total variance. The variables with the greatest weights identified by PCA are attributed to peaks no. 5–8, 11, 13, 14, 16–26 with factor coordinates

either >0.6 or <−0.6, while the remaining peaks (no. 1–4, 9, 10, and 12) were of slightly less importance (smaller weight) with factor coordinates either ⟨0.46;0.58⟩ or ⟨−0.49;−0.59⟩. These PCA results document that the differences between Corsican and non-Corsican honey samples are, although small, not insignificant. Under these conditions, we decided to employ artificial neural networks, which are applicable in situations in which a relationship between the predictor variables (independents, inputs) and predicted variables (dependents, outputs) exists, even when that relationship is very complex [33].

3.2.2. Artificial neural networks

Nowadays, the most common neural network approach to regression-type problems is multilayer perceptrons (MLP) [34,35]. An ANN-MLP based on the back propagation was employed to predict the origin of honey samples based on the pattern of their volatiles. In the first step, the data set was randomly divided by the software into three subsets: (i) *training subset* (1/2 of data), which is used to accomplish the network model training; (ii) *selection subset* (1/4 of data) for checking the network quality within the training process to avoid network overtraining; (iii) *test subset* (1/4 of data) representing the tool to assess the quality of the generated model. The last subset was not used to create the ANN model but was employed for the final validation of the respective ANN model. (*Note:* The subset ratio can be manually changed if higher or lower proportion of each subset is needed.) *Intelligent Problem Solver* was employed for the analysis of data. The search for an appropriate ANN model was restricted only to MLP networks. In the total, 50 networks were tested of which the best ten were retained. The network architecture created for the honey data matrix included an input layer, one hidden layer of neurons, and an output layer. The input layer consisted of 26 neurons (marker compounds), 15–16 neurons in hidden layer (depending on the ANN model), and one neuron in the output layer (origin classification). The ANN was trained using selected parameters from the data sets followed by the validation using an independent data set to estimate the honey origin (Corsica *vs.* non-Corsica). The training started with different initial random weights, and was optimised during the training. Typically, the learning process continues epoch-by-epoch (through single complete training processes) until the synaptic weights and bias level of the network are stabilised [35]. In this study, the network was trained by a back propagation algorithm (100 epochs) followed by a conjugate gradient algorithm (20 epochs). Finally, a
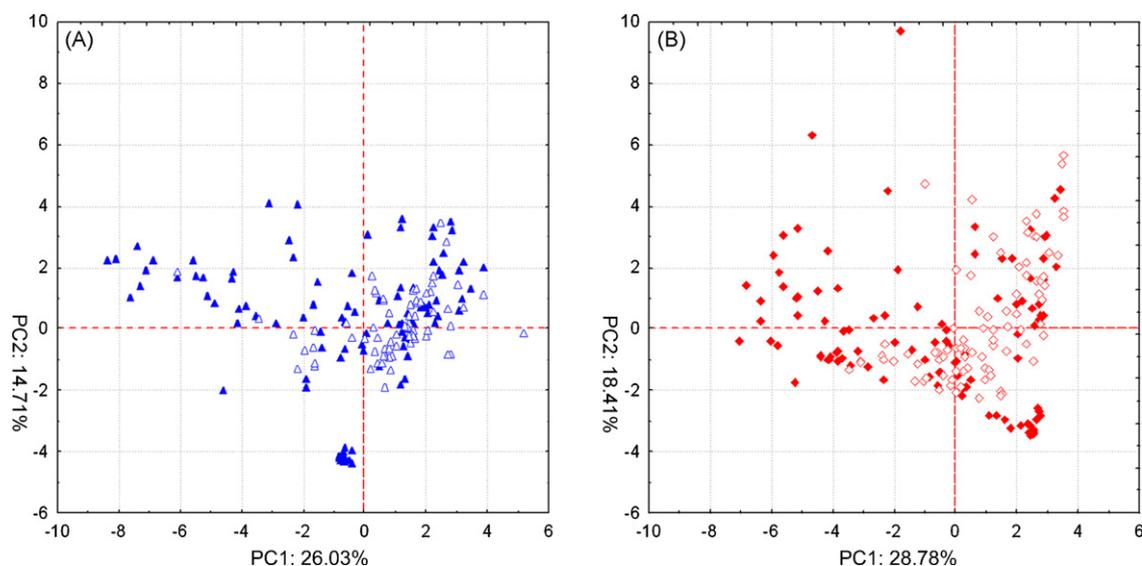

Fig. 2. PCA clustering: (A) year 2006 (▲: Corsica, △: non-Corsica); (B) year 2007 (◆: Corsica, ◇: non-Corsica).

Table 2
Overall summary of ANN-MLP models.

Sampling year	MLP model	Recognition ability [%]		Prediction ability [%]	Classification ability [%]
		Training subset	Selection subset	Test subset	All subsets
2006	26:26–16–1:1	98.9	93.9	93.9	96.2
2007	26:26–15–1:1	100	93.8	91.7	96.4
2006 + 2007	26:26–16–1:1	99.5	92.5	94.6	96.5

network with the smallest error (misclassification of sample origin) was selected. For the separated data sets from the first and second harvest, the MLP network consisted of 26 neurons in the input layer (both harvests), 16 and 15 neurons in the hidden layer for year 2006 and 2007, respectively, and one neuron (both harvests) in the output layer.

The success of the model to classify known objects can be evaluated in different ways: (i) *recognition ability* as the percentage of the samples in the training set (in the case of ANN both training and selection subsets) correctly classified during the modeling step; (ii) *prediction ability* as the percentage of the samples in the test set correctly classified by using the developed model during the training step; (iii) *classification ability* as the percentage of the samples in both training and test sets (in the case of ANN the training, selection, and test subsets) correctly classified by the model [26]. The prediction ability of the models for the first and second harvest was 93.3% and 91.7%, respectively, and the classification ability 96.2% and 96.4%, respectively. The model created for the first harvest (2006) was used to predict samples from the second harvest (2007). The 2006 data were used for formation of respective ANN model with a training, selection, and test (validation) subsets in a ratio of 2:1:1, followed by the testing using the 2007 data as a next validation test set. The prediction ability of this validation set containing 2007 data was somewhat lower (81.3%), probably due to a large variation in the profiles of honey volatiles between these two years possibly caused by different weather conditions in the two harvest years. A low prediction ability (81.9%) was also obtained when the model from the second harvest (2007) was used to predict samples from the first harvest (2006). A more reliable approach seems to be a model that consists of data of a 2-year (and, if possible, even more years) sampling. Under these conditions, each subset contains representative samples from both harvests. Employing this strategy the MLP network consisted of 26 neurons in the input layer, 16 neurons in the hidden layer, and one neuron in the output layer. The prediction and classification abilities of this model were 94.6% and 96.5%, respectively. The overall summary of particular ANN-MLP models is presented in Table 2.

4. Conclusions

In this model study, we used HS-SPME–GC × GC–TOF-MS as a profiling technique in the analysis of honey volatiles, followed by chemometric analysis employing ANN-MLP with the aim to distinguish Corsican honeys (protected denomination of origin region) among honeys harvested in other European countries. Following groups/volatiles (markers) were shown to have the highest discriminant efficiency: (i) *aldehydes* (octanal, nonanal, decanal); (ii) *esters* (ethyl heptanoate, ethyl octanoate, ethyl nonanoate, ethyl decanoate); (iii) *sulphides* (methylsulfanyldisulfanylmethane (dimethyl trisulfide)); (iv) *alcohols* (heptan-1-ol, octan-1-ol, nonan-1-ol, decan-1-ol); (v) *oxygenated aromatics* (benzaldehyde); (vi) *ethers* (lilac aldehyde I, lilac aldehyde II, lilac aldehyde III); (vii) *ketones* (3,5,5-trimethylcyclohex-2-en-1-one (isophorone), 2,6,6-trimethylcyclohex-2-ene-1,4-dione (4-oxoisophorone)).

The results indicate that this approach was successful, fitting to the traceability purpose. The best prediction and classification abilities of the ANN-MLP model were obtained when the data from the two honey harvests were merged. Adding the data from more harvests would probably further improve the performance of this chemometric strategy.

References

[1] Council Directive 2001/110/EC of 20 December 2001 relating to honey, Off. J. Eur. Commun. L10 (2002) 47.
[2] E. Anklam, Food Chem. 63 (1998) 549.
[3] J. Chow, J. Renal Nutr. 12 (2002) 76.
[4] J. Peterson, J. Dwyer, Nutr. Res. 18 (1998) 1995.
[5] R.J. Weston, Food Chem. 71 (2000) 235.
[6] X.-H. Wang, L. Andrae, N.J. Engeseth, J. Agric. Food Chem. 50 (2002) 6923.
[7] Protected Designation of Origin (PDO)/Protected Geographical Indication (PGI), available at http://ec.europa.eu/agriculture/qual/en/1bbab_en.htm via the Internet, accessed July 24, 2008.
[8] G.F. Fairchild, J.P. Nichols, O. Capps Jr., J. Food Distrib. Res. 34 (2003) 38.
[9] Rapid Alert System for Food and Feed (RASFF), Annual Report, 2007, available at http://ec.europa.eu/food/food/rapidalert/report2007_en.pdf via the Internet, accessed September 12, 2008.
[10] CAFIA Annual Report, 2004, available at http://www.szpi.gov.cz/eng/about/article.asp?id=56898&chapter=5&cat=2209&ts=1ec5 via the Internet, accessed September 12, 2008.
[11] S. Bogdanov, K. Ruoff, K.P. Oddo, Apidologie 35 (2004) S4.
[12] B.S. Radovic, M. Careri, A. Mangia, M. Musci, M. Gerboles, E. Anklam, Food Chem. 72 (2001) 511.
[13] T. Cajka, J. Hajslova, J. Cochran, K. Holadova, E. Klimankova, J. Sep. Sci. 30 (2007) 534.
[14] R.M. Pena, J. Barciela, C. Herrero, S. Garcia-Martin, J. Sep. Sci. 27 (2004) 1540.
[15] L. Piasenzotto, L. Gracco, L. Conte, J. Sci. Food Agric. 83 (2003) 1037.
[16] E. Alissandrakis, P.A. Tarantilis, P.C. Harizanis, M. Polissiou, J. Sci. Food Agric. 85 (2005) 91.
[17] E. de la Fuente, I. Martinez-Castro, J. Sanz, J. Sep. Sci. 28 (2005) 1093.
[18] A.C. Soria, I. Martinez-Castro, J. Sanz, J. Sep. Sci. 26 (2003) 793.
[19] J. Dalluge, J. Beens, U.A.Th. Brinkman, J. Chromatogr. A 1000 (2003) 69.
[20] T. Cajka, J. Hajslova, LC GC Eur. 20 (2007) 25.
[21] T. Cajka, J. Hajslova, K. Mastovska, in: S. Otles (Ed.), Handbook of Food Analysis Instruments, CRC Press, Boca Raton, FL, 2008, p. 197.
[22] L. Lebart, A. Morineau, K.M. Warwick, Multivariate Descriptive Analysis. Correspondence Analysis and Related Techniques for Large Matrices, Wiley, New York, 1984.
[23] T. Naes, T. Isaksson, T. Fearn, T. Davies, A User-friendly Guide to Multivariate Calibration and Classification, NIR Publications, Chichester, 2002.
[24] S. Wold, Pattern Recognit. 8 (1976) 127.
[25] H.K.D.H. Bhadeshia, ISIJ Int. 39 (1999) 966.
[26] L.A. Berrueta, R.M. Alonso-Salces, K. Heberger, J. Chromatogr. A 1158 (2007) 196.
[27] http://www.trace.eu.org.
[28] http://www.statsoft.com.
[29] T. Acree, H. Arn, Cornell University, USA, Flavornet and human odor space, http://www.flavornet.org, accessed October 24, 2008.
[30] A.M. El-Sayed, HortResearch, Canterbury Research Centre, New Zealand, The Pherobase, Database of pheromones and semiochemicals, http://www.pherobase.com, accessed October 24, 2008.
[31] B. Massart, Q. Guo, F. Questier, D.L. Massart, C. Boucon, S. de Jong, B.G.M. Vandeginste, Trends Anal. Chem. 20 (2001) 35.
[32] I. Stanimirova, M. Daszykowski, B. Walczak, Talanta 72 (2007) 172.
[33] S. Yi, B.D. Gerardo, Y.-S. Lee, J. Lee, in: M. Gavrilova, et al. (Eds.), Computational Science and Its Applications—ICCSA 2006, Springer, Heidelberg, 2006, p. 144.
[34] T. Olszewski, A. Ryniecki, P. Boniecki, J. Res. Appl. Agric. Eng. 53 (2008) 26.
[35] U. Siripatrawan, B.R. Harte, Anal. Chim. Acta 581 (2007) 63.