Theses of the Ph.D. Dissertation

Zsanett Bodor

Budapest

2022

Hungarian University of Agriculture and Life Sciences

APPLICATION OF CLASSICAL AND CORRELATIVE ANALYTICAL METHODS FOR AUTHENTICATION OF HONEY

Zsanett Bodor

Budapest

2022

PhD School/ Program

Name: Doctoral School of Food Science

Field: Food Science

Head: **Livia Simon Sarkadi, DSc**

Department of Nutrition, Institute of Food Science and Technology, Hungarian University of Agriculture and Life Sciences

Supervisors:

Zoltan Kovacs, PhD

Department of Measurements and Process Control, Institute of Food Science and Technology, Hungarian University of Agriculture and Life Sciences

Csilla Benedek, PhD

Department of Dietetics and Nutrition, Faculty of Health Sciences, Semmelweis University

The applicant met the requirement of the Ph.D. regulations of the Hungarian University of Agriculture and Life Sciences and the thesis is accepted for the defense process.

.....

Head of the Doctoral School

....

Supervisor Supervisor

Table of contents

1 Introduction

Honey is a natural sweetener produced by honeybees *(Apis mellifera)* from nectar, sweet juices of living plant parts, and honeydew. Honey authenticity and its monitoring are in the focus worldwide, which can be explained by its high nutritional and market value. Concerning authenticity, the most important issues are botanical and geographical origin identification and detection of mishandling (addition of exogenous sugar, unallowed processing, heat treatment, etc.). In the last decades, fraudulent activities related to honeys have been increasing. Hence, the origin identification and detection of botanical/geographical mislabeling and other manipulations (including blending with sugar syrups, or overheating) are increasingly important tasks.

Origin identification of honey is usually performed by the combination of three methods: sensory, physicochemical, and pollen analysis. Honey composition and its sensory properties highly depend on its origin, mainly botanical, but the geographical origin is also important. Speculative manipulations can modify the properties of honey as well. All these factors make the detection of honey adulteration quite challenging, involving various techniques. However, on one hand, these methods – such as chromatography techniques, isotopic ratio analysis, mass spectrometry, nuclear magnetic resonance analysis, etc. – are usually time-consuming, and need serious resources, while on the other hand, they are still not entirely accurate. Another problem in the case of botanical origin identification is that there are no compositional criteria for the different unifloral honeys. Some countries define limits for some of the physicochemical parameters of different unifloral honeys, however, these are not available for all the honey types (Thrasyvoulou *et al.*, 2018).

Therefore, there is a demand for methods that can be applied for the aforementioned problems. Moreover, there is also a need for the establishment of a reference database of honeys from different botanical types. This is not available yet, but it would be useful for the authentication and characterization of Hungarian honey. A solution for origin identification and adulteration detection can be a database including results of reference methods combined with the application of correlative techniques, as the latter can provide fingerprint-like data about the analyzed sample. Among these techniques, we can find spectroscopy-based methods such as near infrared spectroscopy (NIR), and instrumental sensory analysis methods (Aykas *et al.*, 2020). As an alternative to the sensory analysis methods, nowadays the use of the artificial senses - such as electronic nose (EN) and tongue (ET) – has started to play a high role in the research in food analysis. According to the literature, ET and NIR were successfully applied for honey analysis aiming at origin identification and adulteration detection (Aouadi *et al.*, 2020). Moreover, the literature available on the application of ET and NIR for the heat treatment detection of honey is missing.

Combined with the reference database, these two methods could serve as alternative tools for compositional and sensory analysis of the honey samples. The use of advanced statistical tools enables analysis of these data together and investigation of application possibilities in the authentication of honey.

2 Objectives

This work aimed to analyze the applicability of reference (moisture content, pH, electrical conductivity pH, ash content, antioxidant properties – total polyphenol content, cupric ion reducing antioxidant power, ferric reducing antioxidant power, and color – L*a*b*, sugars, hydroxymethylfurfural), melissopalynological, and correlative techniques (electronic tongue and near infrared spectroscopy) for the botanical and geographical origin identification, adulteration and heat treatment detection of honeys. Based on these, three main objectives were determined:

1) To apply reference methods, electronic tongue, and near infrared spectroscopy **for the botanical and geographical origin identification**
   a. To give a descriptive characterization of the main Hungarian honey types based on the performed physicochemical determinations
   b. To build botanical and geographical origin classification models for the main Hungarian unifloral honeys using electronic tongue, and near infrared spectroscopy
   c. To investigate the pollen profile of Hungarian honeys from eight main botanical origins
   d. To develop botanical and geographical origin classification models for the main botanical types of Hungarian honeys using near infrared spectroscopy combined with pollen data
   e. To build classification models of the honeys mixed with sugar syrup using acacia and linden honeys
2) To apply mainly near infrared spectroscopy and electronic tongue **for the adulteration detection of honey** supported by reference measurements
   a. To provide descriptive analyzes of honeys mixed (adulterated) with sugar syrup using the main physicochemical methods (pH, electrical conductivity, moisture content)
   b. To develop classification models of the different honey types (acacia, linden, rape, sunflower, honeydew) adulterated with rice, F40 (high fructose corn syrup), and glucose-fructose syrup to see the discrimination efficiency of the electronic tongue and near infrared spectroscopy
   c. To develop predictive models to regress on the added sugar syrup concentration of the aforementioned honey types adulterated with the syrups
3) To apply electronic tongue, near infrared spectroscopy, and reference methods **for the detection of low and high-level heat treatment of honey**
   a. To provide descriptive analyzes of the honeys (acacia, bastard indigo, sunflower) after the application of the heat treatment (40°C, 60°C, 80°C, 100°C for 60, 120, 180, and 240 minutes) using reference methods such as determination of color, pH, electrical conductivity and moisture content)

b. Reveal the efficiency of hydroxymethylfurfural analysis in the detection of heat treatment of honey
c. To develop classification models for the discrimination of temperature, time, and heat treatment level using near infrared spectroscopy and electronic tongue

3 Materials and methods

In this section, the materials and methods and the used sample preparations are going to be introduced. My thesis is separated into three main parts, therefore the materials and methods are subsectioned according to these. The first main part contains the botanical and geographical origin identification studies: namely the basic botanical and geographical origin identification study (BBGOIS), origin identification study extended with pollen analysis (OISWP), and the authenticity study (AUS). The second main part focuses on the sugar syrup adulteration experiments that have two subparts: the sugar syrup adulteration preliminary study (SSAPS), and the sugar syrup adulteration study extended with lower concentrations (SSAWLC). The last part includes of the heat treatment experiment (HTE). The samples and the methods are going to be described in this chapter.

3.1 Materials

3.1.1 Honey samples of the botanical and geographical origin identification

3.1.1.1 Samples of the basic botanical and geographical origin identification study

In the origin identification study, numerous honey types such as 28 acacia (*Robinia pseudoacacia*), 15 linden (*Tilia* spp.), 15 rape *(Brassica napus)*, 11 sweet chestnut (later chestnut - *Castanea sativa)*, 11 milkweed (*Asclepias syriaca*), 17 sunflower (*Helianthus annuus*), 8 bastard indigo (*Amorpha fruticosa*), and 11 multiflora honeys were analyzed. In addition, some rare honey types were also investigated. The samples were collected directly from the beekeepers. The main regions of Hungary were Alföld (Great Plain), Kisalföld (Small Plain), Északi-középhegység (Northern Mountains), Nyugat-magyarországi-peremvidék (Western Hungary), Dunántúli-középhegység (Transdanubian Mountains), and Dunántúli-dombság (Transdanubian Hills) (Dávid, 2013)

3.1.1.2 Samples for origin identification study extended with pollen analysis

In this study, 87 samples were analyzed: 19 acacia (*Robinia pseudoacacia*), 11 linden (*Tilia* spp.), 10 rape *(Brassica napus)*, 10 chestnut *(Castanea sativa)*, 10 milkweed (*Asclepias syriaca*), 10 sunflower (*Helianthus annuus*), and 7 bastard indigo (*Amorpha fruticosa*). All the samples were collected directly from beekeepers from different regions of Hungary. The main regions of Hungary were the same as in 3.1.1.1.

3.1.1.3 Samples for the authenticity study

In this part 12 acacia (*Robinia pseudoacacia*) and 9 linden (*Tilia spp.)* authentic Hungarian honeys were examined, collected from beekepers. In addition three-three mixtures of sugar syrup and acacia or linden honey were prepared in a way to have 10:90, 20:80, 50:50 sugar syrup: honey ratios *(% w/w*). These honeys were labelled as RP10%, RP20%, RP50% for acacia and TI10%, TI20% and TI50% for linden, where the numbers denote the concentration of the syrup.

3.1.2 Sugar syrup adulteration studies

This study part was separated to two parts, however, only the results of the SSAWLC study is shown here, as it is included in the new scientific results.

3.1.2.1 Sugar syrup adulteration study extended with the lower concentrations

In the extended experiment acacia, linden, honeydew, rape, and sunflower honeys were used. The samples were mixed with the syrups at 3%, 5% and 10% ratios. Acacia (RP) and linden (TI) honeys were mixed with rice – (Bio Reis Syrup, dm-drogerie markt GmbH & Co. KG, Karlsruhe, Germany) and K-Sweet F40 – FS – (magas fruktóztartalmú kukoricaszirup, Kall Ingredients, Tiszapüspöki, Hungary). Sunflower (HA), honeydew (HD), and rape (BN) honeys were mixed with the rice, F40, and a self-made glucose/fructose syrup (GF). The GF syrup was prepared with the ratio of 60:40 fructose:glucose: 240 g of analytical grade fructose and 160 g glucose were weighed in a beaker and 100 ml of distilled water was added. All the samples were prepared in three replicates (R1, R2, R3).

3.1.3 Heat treatment experiment

Acacia (RP-*Robinia pseudoacacia*), bastard indigo (AF-*Amorpha fruticosa*), and sunflower (HA-*Helianthus annuus*) honey types were collected directly from beekeepers, making sure that honeys have not been heat treated (control samples) before. Three bottles of 1 kg honey (R1, R2, R3) were used from each type, and the honeys were from the same barrel. Honeys were weighed (50 g each) into 51-51 portions (17 from each bottle of honey) of 100 ml glass sample holders with plastic cap allowing tight closure of the bottle. During the heating no dew condensation was observed. Venticell 111 drying chamber (MMM Medcenter Einrichtungen GmbH., München Germany) was used for the heat treatment of the honeys at 40°C, 60°C, 80°C, or 100°C for 60, 120, 180, or 240 minutes, resulting in 17 levels of heat treatment (time temperature combinations e.g.: 60°C 120 minutes), including control (unheated) and 51 samples per type of honey type (altogether 153 samples).

3.2 Methods

3.2.1 Reference methods

The reference methods such as **moisture content, pH, ash content**, and electrical conductivity were applied according to the method book of the International Honey Commission (Bogdanov, 2009).

Sugar composition (sucrose, fructose, glucose) was determined using RP-HPLC (Waters, Milford, Massachusetts, USA). The device was equipped with a refractive index detector and Kromasil 100-5 NH2 MZ column (250 mm × 4.6 mm, particle size: 5 µm). Throughout the analysis the flow rate was 1.5 ml/minute, and the whole analysis was performed at 25 °C. As mobile phase a solution of water:acetonitrile 28/72*% v/v* was used. All the samples were analyzed in two replicates, where 1 g of honey was dissolved in distilled water and filled up to volume in a 100 ml

volumetric flask. After homogenization the sample solutions were filtered through Chromafil XTRA RC45/24 filter using syringe. For the analysis 10 µl of solution was injected.

For the three **antioxidant properties** the same sample preparation was applied: 10 times *w/v* solution was prepared in a 10 ml volumetric flask. All the parameters were determined using a Thermo Helios Alpha (Thermo Fischer Scientific Inc., Waltham, Massachusetts, United States) UV-VIS spectrophotometer (±0.001 units of absorbance, 1 cm of light path). Folin-Ciocalteu method (Singleton and Rossi, 1965) was applied to determine the total polyphenol content of the honey samples. Ferric reduction antioxidant power of the honey samples was determined according to the method of Benzie & Strain (1996). Cupric ion reducing antioxidant power (CUPRAC) was analyzed according to the method of Apak *et al.* (2008).

Color determination of the honey samples for the origin identification study were analyzed using Konica Minolta 410 colorimeter (Konica Minolta Inc. Chiyoda City, Tokyo, Japan) in five replicates per sample. For the calibration of the device the white tile provided by the producer was used. Color determination for the heat treatment study was done using ColorLite sph850 (ColorLite GmbH, Germany). For the calibration of the device distilled water was used.

3.2.2 Melissopalynology

Acetolysis method was applied for the preparation of the honey samples (Erdtman, 1960). The analyses consisted of the steps described in our study (Bodor *et al.*, 2021)

3.2.3 Sensory analysis of the acacia and linden samples

Sensory analysis of the honey samples has been performed in a sensory laboratory. The analysis and the laboratory were designed to fulfill the requirements of International Organization for Standardization (ISO) standards (ISO, 1994, 2003, 2007). The sensory panel consisted of 12 members, all the samples were analyzed in two sessions. The linden and acacia samples were analyzed separately. The two unifloral types were also analyzed related to different sensory descriptors chosen from the aroma wheel of honey after testing them (Piana *et al.*, 2004).

3.2.4 Rapid correlative techniques

3.2.4.1 Near infrared spectroscopy (NIR)

In my thesis benchtop and handheld devices were used. The benchtop device was a MetriNIR analyzer (MetriNIR Research, Development and Service Co., Budapest, Hungary) that operates in the spectral range of 740-1700 nm, with 2 nm spectral step. The handheld instrument was a NIR-S-G1 (InnoSpectra Co., Hsinchu, Taiwan) that allows spectral acquisition in the range of 900-1700 nm, with 3 nm wavelength step. During all the measurements transflectance setup was applied. The samples were filled in the cuvette ensuring that no bubbles present in the layer. During the MetriNIR measurements the layer thickness was 0.5 mm and the cuvette was temperature controlled at 25°C using circulation around the sample in the coat of the cuvette. The handheld

measurements were performed in a cuvette without temperature control with the layer thickness of 0.4 mm. In each experiment the samples were measured in a randomized order.

3.2.4.2 Electronic tongue

For honey analysis an Alpha MOS ASTREE (AlphaM.O.S., Toulouse, France) potentiometric electronic tongue equipped with an Ag/AgCl reference electrode and seven chemical modified field effect transistor (CHEMFET) electrodes (ZZ, JB, JE, GA, HA, CA, BB) was applied. As a general set up all the samples were recorded through 120 second with 1 second intervals. Then, after every sample cleaning of the sensors was performed for 20 seconds. Before data processing the last 10 seconds of each signal recording were averaged and used for the data analysis. In the case of all the experiment 100 ml of honey solution was prepared using ten times dilution.

3.2.5 Statistical analysis

3.2.5.1 Reference parameters

Descriptive statistics

Descriptive statistics were applied in the case of all the experiments for the reference methods and the sensory profile analysis. The average and the standard deviation, minimum and maximum of the samples were calculated using the Miscrosoft Excel 365 software (Microsoft Corporation, Redmond, Washington, USA).

3.2.5.2 Pollen analysis of the origin identification study extended with pollen analysis

Pollen data was evaluated with the TILIA software where the TILIAGRAPH (Grimm, 1991) was used for the visualization of the results on a pollen spectra. The diagram was created from the taxa that presented in honey in higher than 2%.

Principal component analysis-Linear discriminant (PCA-LDA) analysis

PCA-LDA was applied for the classification of the dataset according to the botanical groups. The model was validated using threefold cross-validation.

Data fusion and models

The data fusion was performed using low-level data fusion approach. During the fusion the pretreated (pretreatment having the best classification accuracy of validation) NIR spectra was concatenated with the pollen spectra (after the exclusion of the pollen <2%) and with the reference (average of the pH, moisture, and ELC per sample) parameters (first level preprocessing). After the fusion the scaling and mean centering of the data was performed (second level preprocessing) (Campos and Reis, 2020). During the scaling the mean centered data points were divided by the standard deviation of the variable. In the case of the geographical origin the main regions of Hungary were class variables (described in 3.1.1.2). The models were validated using threefold cross-validation. Moreover, - as PCA-LDA was performed - PC number optimization was used.

3.2.5.3 Data analysis of the NIR data

Pretreatment and PC number optimization of the NIR data

For all the analysis types a pretreatment optimization was performed, where in total 41 different pretreatment types were tried. Amongst pretreatments Savitzky-Golay smoothing, multiplicative scatter correction (MSC), standard normal variation (SNV), detrending (deTr), and their combinations were applied. Moreover, in the case of all the model PC number optimization was performed.

Principal component analysis-Linear discriminant analysis (PCA-LDA)

In the case of all the PCA-LDA models in all of the experiments the models were built trying all the pretreatment combinations and models having the best average validation accuracy were chosen, moreover the used principal component numbers (PCs) used for the model was also optimized, choosing the model with the highest validation average accuracy. Classification models were built separately for the different botanical groups in the case of the **adulteration experiments (SSAWLC)**. Moreover, within the botanical groups, models were built using data of all the syrup types (with all their respective samples) and then separately for the data of the different syrups using three-times cross-validation.

PCA-Linear discriminant analysis was also used in the **heat treatment experiment (HTE)**, models were built for the classification of the applied temperature level (control, 40°C, 60°C, 80°C, 100°C), the applied time interval (control, 60, 120, 180, 240 minutes) and also for the heat treatment level (temperature time combinations resulting in 17 levels including the control). The models were validated using threefold cross validation.

3.2.5.4 Data analysis of the electronic tongue data

Before the analysis by the electronic tongue, the sensory data were drift-corrected using the "additive correction relative to reference samples" (ACRRS) (Kovacs *et al.*, 2020).

Linear Discriminant analysis of the ET data

In the case of the **basic botanical and geographical origin identification study (BBGOIS)** models were firstly built for the classification of the main botanical groups (acacia, rape, bastard indigo, sweet chestnut, linden, sunflower, honeydew, milkweed and multiflora). As a second phase to be able to make comparison with the pollen extended study the six groups (acacia, rape, bastard indigo, sweet chestnut, linden, sunflower) were used for classification model of botanical origin discrimination. After this step all the six groups were used for the classification model of the geographical regions. However, due to the higher effect of the botanical origin, the models were built separately for all the honey types. In this case the geographical origin classifications were built for the different counties, with the exception of the bastard indigo, where the districts were the class variables (there were only two counties in their case). All the models were built using threefold cross-validation.

In the case of the **authenticity study (AUS)**, the models were built separately for the two types of honey for the classification of the authentic acacia or linden honeys and the mixture of the honeys with the sugar syrups in 10%, 20% and 50%. The models were validated using threefold cross-validation.

In the case of the **sugar syrup adulteration experiment (SSAWLC)** the models were built for the classification of the different syrup mixed sunflower honeys. Four models were built: one that contained the control and all the syrup mixtures (10 levels) and three models for the three different syrup mixtures (F40, rice, GF) having 4-4 groups per models. All the models were validated with threefold cross-validation.

In the case of the **heat treatment experiment (HTE),** models were built for the heat treatment levels, temperature groups and time intervals, separately for the three types of honey (acacia, bastard indigo, sunflower). The models were validated using threefold cross-validation after the pretreatment and outlier detection.

4 Results and discussion

4.1 Results of the origin identification study

In this paragraph the results of the studies related to the botanical and geographical origin identification will be presented.

4.1.1 Results of the basic botanical and geographical origin identification study

4.1.1.1 Descriptive analysis of the Hungarian unifloral, honeydew and multiflora honeys

In this section the descriptive sheets of the unifloral honeys collected from different regions of Hungary are going to be presented. Aim of these descriptive sheets was to summarize the physicochemical properties of the Hungarian unifloral honeys. Along with the basic botanical and geographical origin identification study the descriptive tables of eight unifloral honeys (acacia, rape, bastard indigo, milkweed, sweet chestnut, linden, sunflower, honeydew) and multiflora honeys were prepared including the aforementioned reference methods and the melissopalynology.

4.1.1.2 Botanical and geographical origin identification based on electronic tongue data

The origin identification model built for the classification using linear discriminant analysis for the acacia, linden, sweet chestnut, honeydew, sunflower, rape, bastard indigo, milkweed, and multiflora honeys provided average recognition and prediction abilities of 57.07% and 57.09%, respectively after the threefold cross-validation. Owing to these weak results, another model was built for the classification of the six main groups (acacia, linden, sunflower, rape, chestnut, and bastard indigo). The LDA classification model provided the average classification accuracies of 70.91% and 70.51% during the training and cross validation. Chestnut honey was classified correctly, while acacia showed correct classification in 93.75% showing misclassification belonging to the bastard indigo (5.92%) and rape (0.33%) after validation. The bastard indigo honey was misclassified also belonging to the acacia (32.83%) and rape (1.5%). The rape honey showed a strong misclassification, partly belonging to the acacia (70.15%), so we could not separate these two groups from each other. The sunflower honey was classified correctly in 86.21%, misclassification was found for samples belonging to the linden group (13.79%). The linden honey was classified correctly in 50%, misclassified samples belonging to all the other groups in the range of 22.98-1.72%, except for bastard indigo.

The results of the geographical origin classification using the data of six aforementioned groups built for the classification of the main geographical regions of Hungary provided average training and validation accuracies of ~59%. These low accuracies can be due to the fact that the effect of the botanical origin is higher than those of geographical origin. Therefore, if the samples are from the same plant, but different geographical regions, this can lead to misclassifications. Because of the higher effect of the botanical origin, the models will be demonstrated separately for the different botanical types. Classification models were built for the discrimination of the counties or

districts (in the case of bastard indigo). The models for sunflower, bastard indigo, and linden provided 100% correct classification for all the groups. The model of the chestnut honey showed correct classification of the groups during the training and 98.17% during validation. In the case of the rape honey, the average training and cross-validation accuracies were 97.24%, 90.75%, respectively. The model of acacia was a somewhat less accurate, where the average training and validation were 71.85% and 70.68%.

4.1.2 Results of the origin identification study extended with pollen analysis

4.1.2.1 Results of the pollen analysis

Melissopalynological analysis of the 87 honey samples from the eight botanical origins provided 107 identified taxa in the samples.

Based on the melissopalynological data a reference database was built up containing the most abundant pollen (>2%) types in the analyzed honeys. PCA-LDA models built for the classification of the botanical origin showed that the average training and validation accuracies were 99.14% and 90.22%, respectively. The classification of the validation dataset showed that the rape and the chestnut groups were classified correctly. The acacia samples showed again misclassification as belonging to the bastard indigo honey group in 5.5%. The bastard indigo samples also showed a misclassification as belonging to the rape honeys in 14.16%. Moreover, 9.02% of the linden group was classified as sunflower honey. The PCA-LDA model of the geographical origin classification provided weak results of 68.57% and 51.93% of average classification accuracy during the training and the validation, respectively.

4.1.2.2 Results of the fusion of NIR, physicochemical and pollen data

The fusion of the data of the reference methods, pollen and NIR provided improved accuracies both for the classification of the botanical and geographical origin. The fusion of the three datasets provided higher than 99% accuracy, even after the threefold cross validation in both botanical and geographical origin classification. In one case misclassification was found, where the acacia sample group showed misclassification as belonging to the bastard indigo group in 4.20% and 4.21% during the training and cross validation. The geographical origin classification model provided 100% accuracy during training and validation.

4.1.3 Results of the authenticity study highlighting the sensory results

The results of the reference methods showed that the pH, electrical conductivity, moisture content, total polyphenol content, CUPRAC, and FRAP showed a decreasing tendency with the higher syrup concentration. In the case of all the parameters it can be seen that the two honey types (acacia and linden) had different results regarding the discrimination of the authentic and the mixture samples.

This was also valid for the sensory profile analysis. In the case of the **acacia** only four parameters showed that there is a significant difference among the authentic honey and the syrup

blends. The RP10% and RP50% showed significantly less intense sweet and flowery taste, while RP10% had significantly more intense fruity odor. The honeys containing 20% and 50% syrup were characterized by significantly higher caramel taste intensity.

In comparison with pure acacia, the **linden** honeys mixed with differed significantly in eight parameters from the reference sample. The honey containing 10% (TI10%) syrup was characterized by significantly stronger fresh odor, while TI20% was significantly sourer compared to the reference sample. Moreover, TI50% was scored with significantly less intense bitter and medicinal flavor. All the adulterated samples had significantly less intense sweet taste, global odor intensity and global taste intensity, and aftertaste persistence.

Results of the **electronic tongue** seemed to be promising in the differentiation of the mixtures from the authentic samples. In the case of the **acacia** the validation accuracy was 99.22%, where the adulterants were all classified correctly but the control honey showed misclassification in 3.11% belonging to the group of the RP10%. In the case of the **linden** honey lower average training and validation accuracies were obtained, however in this case the authentic honey was classified correctly.

4.2 Results of the extended sugar syrup adulteration study with low concentrations

4.2.1 Results of PCA-LDA models of the NIR dataset

PCA-Linear discriminant analysis models of the NIR of the different botanical groups showed correct classification of the control sample in all of the models. The average training and validation accuracies were also 100% in most of the cases after the model optimization.

4.2.2 Results of the LDA models of the electronic tongue data of sunflower honey and its syrup mixtures

The linear discriminant analysis model built for the classification of all the groups of the **sunflower** and its mixtures provided the average training and validation accuracies of 91.13% and 37.77%, respectively after the threefold cross-validation. During the training the control sample was classified correctly and none of the samples showed overlapping with the control. However, in the validation the control showed misclassification as belonging to the 3% rice syrup-containing samples in 11%. During the validation also none of the samples were classified as a control, showing that the control could be separated from the adulterated samples. The model of the **F40 syrup** adulterated honeys provided the average recognition and prediction accuracies of 100% and 75%. Throughout the validation the control showed misclassification as belonging to the honey that contained 3% of the syrup in 11%. The **rice syrup** model also provided the average training accuracy of 100%, and better validation accuracy of 88.92%. In the validation dataset the control showed misclassification as belonging to the 3% adulterant in 11%. Similarly to the results of NIR the **GF syrup** detection was weaker than the F40 and rice syrup in the case of the sunflower honey.

The average training and validation accuracies were 100% and 63.92%. During the validation misclassification were found except the control, which was classified correctly.

4.3 Results of heat treatment study with special regard to the NIR and electronic tongue

4.3.1 Results of the near infrared spectroscopy

Sunflower honey:

PCA-LDA model built for the classification of the temperature, time and treatment level provided the average classification (validation) accuracy of 84.01%, 62.83% and 80.81%, respectively.

Temperature classification model showed that the control honey was classified correctly even after the validation, however the 40°C-group showed less than 3% misclassification as belonging to the control. The higher temperature groups (60°C, 80°C, 100°C) could be completely separated from the control honey, but showed misclassifications as belonging to each other.

The model of the **time intervals** showed correct classification of the control, but the 60 minutes treated sample group showed misclassification as belonging to the control in 2.50% and 5.01%, during the training and validation, respectively. The higher time intervals did not overlap with the control.

The classification model of the **heat treatment level** showed that the control was classified correctly during both training and validation, moreover, no misclassification was found belonging to the control. At higher levels, especially above 60°C and 60 minutes, there were misclassifications.

Bastard indigo

The models of the bastard indigo were slightly different from the models of the sunflower honey. The temperature model showed average validation accuracy of 84.93%, the time model provided results of 60.93% and for the level 74.93% average validation accuracy was obtained.

The model built for the classification of the **temperature levels** showed 100% correct classification of the control honey, and none of the groups showed misclassification as belonging to the control

The **time interval** classification was weaker, however the control honey showed 100% correct classification, and all the other honeys showed misclassification as belonging to each other, but none of the groups were misclassified belonging to the control.

The classification model of the **heat treatment levels** provided 100% correct classification of the control honey and none of the other treatment groups were misclassified as belonging to it.

4.3.2 Results of the electronic tongue analysis

Sunflower

LDA model built for the classification of the temperature, time and treatment level provided the average classification accuracy after the validation of 84.28%, 54.10% and 67.20%, respectively

The model built for the **temperature** classification showed that the control honey was classified correctly in 96.46% and 85.65% during the training and validation, where misclassification was found as belonging to the 40°C. The higher temperature groups could be completely separated from the control honey.

The model of the **time intervals** showed correct classification of the control, however all the other groups showed misclassifications as belonging to each other, therefore the effect of the temperature was higher.

The classification of the **heat treatment level** group showed that the control was classified correctly during the training but the through the validation misclassification was found in 7.08% as belonging to the group heated at 40 °C for 180 minutes sample.

Bastard indigo

The models of the bastard indigo were different from the models of the sunflower honey, where the temperature model showed average validation accuracy of 75.57%, the time model provided worse results of 30.29% and for the level 54.43% average validation accuracy was obtained.

Model of the **temperature** classification showed that after the validation the control was classified correctly in 85.65%, and the misclassification was belonging to the 40°C group. The other honeys did not show misclassification as belonging to the control with the exception of 40°C where the misclassification was 9.28% during the validation.

The **time level** classification was very weak, where the control honey showed 56.17% (misclassified to 60 minutes). The other time groups also showed misclassification as belonging to the control during the validation, where the 60 minute samples showed 15.54%, the 120 2.20% and the 180 minutes 2.36% of misclassification.

The model of the **heat treatment levels** provided the 85.65% of correct classification of the control, the misclassification was found as belonging to the honey heated at 40°C for 60 minutes. Moreover these samples showed also misclassification as belonging to the control in 8.25% throughout the validation.

These results show also that only the 40°C (allowed) heat treatment level could not be discriminated from the control, this showing the power of electronic tongue in the detection of the heat treatment.

5 Conclusion and recommendations

The thesis focused on the origin identification, adulteration and heat treatment detection of Hungarian honey using reference parameters, melissopalynology, and correlative methods.

Along the basic botanical and geographical origin identification study the descriptive tables of eight unifloral honeys (acacia, rape, bastard indigo, milkweed, sweet chestnut, linden, sunflower, honeydew) and multiflora honeys were prepared including the aforementioned reference methods and the melissopalynology. In the future the expansion of this database would be useful to have continuous data from the different years.

An AlphaASTREE potentiometric electronic tongue was used for the botanical and geographical origin identification of the aforementioned honey types. When all the nine honey types were analyzed together, the average prediction and recognition ability of the LDA model was 57.09% after threefold cross-validation. The high scatter of the honeydew and multiflora honeys could cause this weak classification accuracy. When these honey types left out from the model the classification accuracy was 70.51%. The chestnut honey was classified correctly, while the acacia, rape, and bastard indigo honeys grouped together, providing misclassification belonging to each other. The LDA models using the ET data built for the classification geographical regions provided ~59% validation accuracy. The reason of the weak model could be the higher effect of the botanical origin. Therefore, the models were analyzed separately for the six honey types and the counties or districts were used as group variables. These models provided better classification accuracies ranging from 70.68%-100% average classification accuracy.

The origin identification study extended with pollen analysis showed that the pollen spectra of the honeys contribute highly to the separation of the botanical groups, where >90% classification accuracy was obtained for the acacia, bastard indigo, rape, chestnut, sunflower, and linden honeys. The fusion model of the NIR, melissopalynology and physicochemical properties (pH, EC, moisture) provided higher than 99% classification accuracy. From this we can conclude that the combination of these techniques was reliable.

In the authenticity study authentic acacia, linden honeys, their blends with syrup in 10%, 20%, and 50%, were analyzed using the reference methods and electronic tongue and sensory profile analysis. The results showed that in the case of the sensory profile analysis the panel could discriminate the adulterated acacia honeys from the authentic in four parameters, while in the case of the linden eight parameters showed significant difference. Moreover, the electronic tongue showed that the control sample could be separated completely from the adulterated samples in the case of linden honey. In the case of acacia, the pure honey showed misclassification belonging to the 10% adulterated honey in 3.11%.

As an extension of this sunflower, linden, acacia, rape and honeydew honeys were mixed with the high fructose content sugar syrup (F40), rice syrup, and glucose-fructose (self-made) syrup in 3%, 5% and 10%. The samples were analyzed using the benchtop NIR instrument and the sunflower honeys were analyzed with the electronic tongue as well. The electronic tongue results showed that during the validation the LDA model did not provide good accuracy when analyzing all the syrups in one model, however if the models were built separately for the syrups, where the control honey was only misclassified with the 3% adulterated group. In the sugar syrup adulteration study extended with lower concentrations acacia, the PCA-LDA models provided in all the cases higher than 94% classification accuracy of the total sample set, but 100% classification accuracy of control. In the future it would be useful to expand this database with honeys from different regions from the same type using the syrups.

The heat treatment experiment was performed on the acacia, bastard indigo, and sunflower honeys, where the honeys were heated at 40, 60, 80, and 100 °C for 60, 120, 180, and 240 minutes. The samples were analyzed using reference methods such as moisture, pH, electrical conductivity, color and HMF content, and the correlative NIR and electronic tongue. NIR was more accurate comparing with the electronic tongue. Another important notice is that the models were better in the case of the sunflower, bastard indigo honeys comparing with the acacia. The reason for this could be that the acacia is less rich in useful components such as minerals, antioxidants, moreover, the sensory characteristics and aroma of this honey are also weaker. Therefore, these results were excluded from the new scientific results.

In summary it can concluded that both NIR and electronic tongue could be efficient tools in the origin identification, adulteration and heat treatment detection. In the future it would be useful to analyze the samples with both instruments and use these as a fused dataset.

6 New scientific results

For the purpose of these new scientific findings, benchtop spectrophotometer refers to MetriNIR (MetriNIR, Research Development and Service Co., Budapest, Hungary), whereas handheld spectrophotometer refers to NIR-S-G1 (InnoSpectra Co., Hsinchu, Taiwan). E-tongue refers to the Alpha MOS ASTREE potentiometric electronic tongue (AlphaM.O.S, Toulouse, France) equipped with seven sensors developed for food application (BB, HA, ZZ, GA, CA, JE, JB), a reference electrode and a 16-position autosampler. Pollen spectra obtained with melissopalynological using acetolysis method. Sensory profile analysis was performed according to the requirements of International Organization for Standardization (ISO) standards (ISO, 1994, 2003, 2007).

6.1 New scientific findings focusing on botanical and geographical origin identification

1) Reference database was established containing physicochemical parameters (moisture, pH, electrical conductivity, ash content, color (L*a*b*), antioxidant parameters) of 137 Hungarian honey from the most common nine botanical types (acacia, sunflower, linden, chestnut, milkweed, honeydew, rape, bastard indigo, and multiflora) originating from all regions of Hungary (Északi-középhegység, Dunántúli-középhegység, Alföld, Kisalföld, Dunántúli-dombság and Nyugat-magyarországi-peremvidék), and collected between 2015-2020. This database can be used as reference for the authentication of common Hungarian honey types.

2) Pollen spectra based database of 87 Hungarian honey from the most common eight botanical types (acacia, sunflower, linden, chestnut, milkweed, honeydew, rape, bastard indigo) from all regions of Hungary (Északi-középhegység, Dunántúli-középhegység, Alföld, Kisalföld, Dunántúli-dombság and Nyugat-magyarországi-peremvidék), collected between 2015-2020 was established.

3) PCA-LDA models were built for the first time using the low-level data fusion of physicochemical data (pH, moisture, electrical conductivity) pollen spectra, and NIR spectra for the botanical and geographical origin identification of Hungarian authentic honeys from eight botanical origin (acacia, linden, sunflower, chestnut, honeydew, milkweed, sunflower, bastard indigo), and all regions of Hungary (Északi-középhegység, Dunántúli-középhegység, Alföld, Kisalföld, Dunántúli-dombság and Nyugat-magyarországi-peremvidék) collected between 2015 and 2020. The models provided high classification accuracies for botanical and geographical origin identification (99.30% for botanical origin, 100% for geographical origin identification in cross validation).

4) Electronic tongue was used for the first time for botanical and geographical origin identification of 50 Hungarian authentic honeys from nine botanical origins (acacia, linden, sunflower, chestnut, honeydew, milkweed, sunflower, bastard indigo, multiflora), and from four regions of Hungary (Északi-középhegység, Alföld, Dunántúli-dombság and Nyugat-magyarországi-

peremvidék) collected between 2012 and 2016 (mostly from 2015). LDA models provided better classification accuracies for botanical origin identification (70.91% and 70.51% in training and validation, respectively, after excluding the data of honeydew, milkweed and multiflora honeys), than for the geographical origin identification (59.46% and 59.40% in training and validation accuracy, respectively). Improved accuracies were obtained for the geographical origin identification analyzing the botanical groups separately (training accuracies ranged between 71.85 and 100%, and validation accuracies between 70.68 and 100%).

6.2 New scientific findings focusing on sugar syrup adulteration detection

5) Sensory profile analysis was applied on Hungarian authentic honeys (acacia and linden honeys collected in 2016 from Heves and Pest counties) and their blends with sugar syrup (honeys mixed with sugar syrup at 10%, 20% and 50%). Significant differences were found between the authentic honeys and their blends with 10% sugar syrup in three (fruity odor, sweet and flowery taste), and in five (odor and taste intensity, fresh odor, sweet taste, and aftertaste persistence) sensory parameters for acacia and linden honeys, respectively.

6) Electronic tongue was used for the first time for sugar syrup adulteration detection of Hungarian authentic sunflower honey mixed with different sugar syrups at different levels (rice syrup, F40 – high fructose content sugar syrup, and self-made glucose-fructose syrup (80% of 40/60 glucose/fructose +20% water) each applied at 3%, 5% and 10%, separately). Electronic tongue combined with LDA was able to discriminate the adulterated honeys from the authentic honeys with 100% accuracy and provided misclassification of the authentic honey belonging to samples containing 3% sugar sirup (in 11% belonging to 3% rice sirup and F40 sirup, respectively).

7) Benchtop spectrophotometer was used for the first time for sugar syrup adulteration detection of Hungarian authentic honeys from five botanical origins (acacia, linden, sunflower, rape, honeydew) mixed with different sugar syrups at different concentrations (each honey was mixed with rice and F40 syrups, and rape, sunflower and honeydew honeys were mixed with self-made glucose-fructose syrup (80% of 40/60 glucose/fructose +20% water) in 3%, 5% and 10%, respectively). PCA-LDA models of all the honey types for all the different model variations (models including all the syrups, or syrups separately) provided the complete discrimination of the adulterated and pure honeys (100% classification accuracy was obtained for the control in all the cases). In this regard honeys could be clearly discriminated from the 3% mixtures.

6.3 New scientific findings of the heat treatment study

8) Handheld near infrared spectrophotometer was applied for the first time for the detection of heat treatment (heated at 40°C, 60°C, 80°C, or 100°C, and at each hold for 60, 120, 180, or 240 minutes) of Hungarian authentic honeys (sunflower, bastard indigo honeys). PCA-LDA models

of the classification of heat treatment levels, temperatures and time intervals provided correct classification of the authentic honeys in the case of both honey types.

9) Electronic tongue was applied for the first time for the detection of heat treatment (heating temperatures: 40°C, 60°C, 80°C, or 100°C, time intervals for each temperature: 60, 120, 180, or 240 minutes) of Hungarian authentic honey samples (sunflower, bastard indigo). PCA-LDA models were built for the classification of temperature level, time interval, and heat treatment levels. In the case of sunflower the control was classified correctly in 85.65%, 100% and 92.92% of the temperature, time and heat treatment level model, respectively. Misclassification of the control was found as belonging to 40°C group (temperature model) and to the honey treated at 40°C for 180 minutes (heat treatment level model). In the case of bastard indigo the PCA-LDA models showed that the control was classified correctly in 85.65% (misclassified to 40°C) in the temperature model, while in time-interval model only in 56.17% (misclassified to 60 min), and the heat treatment level model in 85.65% (misclassified to 40°C 60 minutes honeys). Both honey types showed the complete separation of control from the honeys heated at 60°C or higher temperatures.

7 List of journal publications in the field of studies

Bodor, Z., Benedek, C., Kaszab, T., Zaukuu, J.-L.Z., Kertész, I., Kovacs, Z., Zinia Zaukuu, J.-L., Kertész, I., Kovacs, Z., **2019.** Classical and correlative analytical methods for origin identification of Hungarian honeys. ***Acta Alimentaria*** 48, 477–487. https://doi.org/10.1556/066.2019.48.4.9 ***Q3 - IF 0.650***

Bodor, Z., Ghdir, C., Zaukuu, J.-L.Z., Benedek, C., Kovacs, Z., **2019**. Detection of heat treatment of honey with near infrared spectroscopy. ***Hungarian Agricultural Engineering Sciences.*** 36, 57–62. https://doi.org/10.17676/HAE.2019.36.57

Bodor, Z., Kovacs, Z., Rashed, M.S., Kókai, Z., Dalmadi, I., Benedek, C., **2020**. Sensory and physicochemical evaluation of acacia and linden honey adulterated with sugar syrup. ***Sensors*** 20, 1–20. https://doi.org/10.3390/s20174845 ***Q2 - IF 3.576***

Bodor, Z., Kovacs, Z., Benedek, C., Hitka, G., Behling, H., **2021.** Origin Identification of Hungarian Honey Using Melissopalynology, Physicochemical Analysis, and Near Infrared Spectroscopy. ***Molecules***. https://doi.org/10.3390/molecules26237274 ***Q1 - IF 4.412***

Bodor, Z., Benedek, C., Aouadi, B., Zsom-Muha, V., & Kovacs, Z. (2022). Revealing the Effect of Heat Treatment on the Spectral Pattern of Unifloral Honeys Using Aquaphotomics. ***Molecules,***
https://doi.org/10.3390/molecules27030780 ***Q1 - IF 4.412***