European Food Research and Technology
https://doi.org/10.1007/s00217-023-04347-1

ORIGINAL PAPER

Application of UV–Vis spectroscopy for the detection of adulteration in Mediterranean honeys

Dafni Dimakopoulou-Papazoglou[1] · Nikolaos Ploskas[2] · Salud Serrano[3] · Carolina Santos Silva[4,5] · Vasilis Valdramidis[4,6] · Konstantinos Koutsoumanis[1] · Eugenios Katsanidis[1]

Received: 31 May 2023 / Revised: 24 July 2023 / Accepted: 29 July 2023

Abstract
The present study aimed to identify adulteration of honey with sugar syrups and colorants using UV–Vis spectroscopy, combined with multivariate statistical analysis. A total of 209 honeys were used, including 151 commercial honey samples (thyme, pine, and polyfloral honeys) collected from different countries of Mediterranean (Greece, Malta, Spain, Tunisia, and Turkey) and 58 adulterated Greek thyme honey samples by adding syrups and colorants. Honey adulteration was identified using Principal Component Analysis (PCA) along with Random Forest (RF), Partial Least Squares – Discriminant Analysis (PLS-DA), and Data Driven-Soft Independent Modelling of Class Analogies (DD-SIMCA) using the spectral range of 220–550 nm. Comparatively, DD-SIMCA models produced better results in terms of accuracy and sensitivity in most cases evaluated. The results support the good predictive capability of UV–Vis spectroscopy combined with chemometrics for the determination of honey adulteration, and thus, it could be utilized as a rapid, inexpensive, and simple method.

Keywords Mediterranean honey · UV–Vis spectroscopy · Adulteration · Sugars · Colorants · Multivariate statistical analysis

Introduction

Honey has been characterized as a functional food and its consumption has gained in popularity due to its nutritional qualities and beneficial health effects [1]. Generally, honey is produced by *Apis mellifera* bees from the nectar of flowers and secretion of plants or insects which are collected from a great variety of florals [2]. Honey is composed of 70–80% saccharides (mainly glucose and fructose and in minor quantities maltose and sucrose), 15–20% water and small amounts of beneficial substances, such as proteins, amino acids, phenolics, vitamins, pigments, enzymes, and other biological compounds [3, 4]. The variety of these compounds makes honey a nutritionally valuable product that differs significantly from simple sweeteners like sugar syrups [5] and therefore has a higher commercial value. Therefore, honey is a target for adulteration worldwide [6], usually by the addition of inexpensive ingredients, such as sugar syrups and colorants, an illegal practice that decreases the cost and the nutritional value of honey.

One of the most crucial challenges in quality assurance and food safety is food authentication, which is of concern to regulatory agencies, food producers, distributors, retailers,

✉ Dafni Dimakopoulou-Papazoglou
ddimakop@agro.auth.gr

1 Department of Food Science and Technology, School of Agriculture, Faculty of Agriculture, Forestry and Natural Environment, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece

2 Department of Electrical and Computer Engineering, Faculty of Engineering, University of Western Macedonia, 50100 Kozani, Greece

3 Department of Food Science and Technology, Campus of Rabanales, University of Córdoba, Charles Darwin Building Annex, 14014 Córdoba, Spain

4 Department of Food Sciences and Nutrition, Faculty of Health Sciences, University of Malta, Msida 2080, Malta

5 VTT Technical Research Centre of Finland Ltd, Tekniikantie 1, 02150 Espoo, Finland

6 Department of Chemistry, National and Kapodistrian University of Athens, 15771 Athens, Greece

Published online: 09 August 2023

and consumer advocacy organizations [7]. Detecting honey adulteration can be challenging due to honey's high natural diversity that results from the habitat and the plethora of plant species from which honey is produced, as well as from the maturity level and the processing and storage conditions [7–9]. The most common method for honey adulteration is the addition of low-cost sugars at some point during production or processing [10]. Specifically, this is accomplished by either overfeeding the bees with sugar during nectar flow period to get more honey, or by adding sugar syrups to honey after production [11]. For the adulteration of honey, several low-cost sugars, such as sucrose syrup, glucose syrup, fructose syrup, maltose syrup, corn syrup, cane sugar, rice syrup, beet syrup, inverted syrup, etc., have been used [12–15]. Additionally, another type of adulteration, which occurs to a lesser extent, is the addition of colorants to honey (caramel color, classes E150c and E150d), as it is widely accepted that darker coloured honey is associated with stronger antioxidant activity due to its high phenolic compounds content [3, 16].

The authenticity of honey is crucial for both commercial and health reasons. So far, physicochemical analysis has been successfully used to determine various adulterants such as water, sucrose, invert sugar, dextrin, and starch [17]. However, often classical physicochemical analyses are not enough to determine whether or not the honey is fraudulent, so more sophisticated analytical techniques need to be used, such as high-performance liquid chromatography (HPLC) [18, 19], gas chromatography–mass spectrometry (GC–MS) [20], nuclear magnetic resonance (NMR) spectroscopy [21, 22], and carbon isotope ratio analysis [23, 24]. Despite the efficacy of these techniques, they are time consuming, require expensive equipment, intricate sample preparation, and experienced personnel. Hence, due to the need for rapid results and economic and simple methods for the detection of adulteration in honey, alternative techniques were effectively developed, such as spectroscopic techniques including Fourier transform infrared spectroscopy (FTIR) [5, 25, 26], near infrared spectroscopy (NIR) [13, 27, 28], Raman spectroscopy [29–31], and fluorescence spectroscopy [32, 33].

A promising spectral technique is ultraviolet – visible (UV–Vis) spectroscopy that has been successfully used to determine the botanical and geographical origin of honey [34–38]. However, studies regarding the detection of honey adulteration using UV–Vis spectroscopy are scarce [39–42] and limited only to the adulteration with sugar syrups. Specifically, de Souza et al. [39] utilized one-class classification methods for detecting honey adulteration with sugar syrups (corn syrup, agave syrup, and sugarcane molasses) using 235 samples. Suhandy et al. [40] used Principal Component Analysis (PCA) to discriminate pure (Sumbawa monofloral honey) and adulterated honey with high fructose corn syrup using a dataset of 50 samples. Mitra et al. [41] applied four machine learning classifiers to detect honey adulteration using 19 samples. Valinger et al. [42] utilized Partial Least Squares and Artificial Neural Network models for the detection of adulteration using acacia honey samples (15 pure and 135 adulterated honey samples). To the best of our knowledge, there are no studies using UV–Vis spectroscopy for the detection of honey adulteration that: (i) includes samples from various countries, (ii) aims to detect both colorant and sugar syrup adulterations, and (iii) compares the results of one-class and binary-class classification methods. In view of the above, the objective of the present study was to investigate the suitability of UV–Vis spectroscopy combined with multivariate statistical analysis to detect the adulteration of Mediterranean thyme honeys with sugar syrups, colorants, or the combination of both adulterants.

Materials and methods

Honey samples and sample preparation

A total of 209 honey samples were used in this study including 151 commercial samples (72 thyme, 20 pine, and 59 polyfloral honeys) from different Mediterranean countries (46 honeys from Greece, 42 from Tunisia, 31 from Turkey, 22 from Spain, and 10 from Malta) and 58 adulterated samples. Four commercial thyme honeys collected from Greece (two light-coloured and two dark-coloured honeys), were used in order to prepare the adulterated samples. The adulteration of the samples was performed by adding (a) different sugar syrups, (b) different colorants, and (c) both sugar syrup and colorants. The sugar syrups used were either (a.1) rice syrup (La Finestra sul Cielo, S.A., Italy; ingredients: water, rice 35%) or (a.2) invert sugar (invert sugar, Innovative Cooking S.L., Madrid, Spain; Ingredients: invert sugar syrup, potassium sorbate). The colorants used were either (b.1) ammonia caramel, E150c (Krendal, S.A.S., Italy) or (b.2) sulfite ammonia caramel E150d (Krendal S.A.S., Italy). The honey samples were adulterated by syrups in a concentration level of 5, 10 and 25% (w/w), while the colorants were added either directly to the honey at concentration level of 50, 100 and 150 ppm, or to the rice syrup at the same concentrations and, subsequently to the honey in a concentration level of 25% (w/w). The addition of colorants occurred only in the light-coloured honey samples, as this type of adulteration takes place in order to make the colour of the honeys darker. The preparation and the composition of the adulterated honey samples is shown in Table 1. After preparation, the samples were kept in clean plastic containers and stored under refrigeration until analysis.

Table 1 Composition of adulterated honey samples

Fraud scenario	Sugar addition (w/w) %	Colorant addition
Sugar adulteration	Rice syrup – 5	
	Invert sugar – 5	
	Rice syrup – 10	
	Invert sugar – 10	
	Rice syrup – 25	
	Invert sugar – 25	
Colorant adulteration		E150c – 50 ppm
		E150c – 100 ppm
		E150c – 150 ppm
		E150d – 50 ppm
		E150d – 100 ppm
		E150d – 150 ppm
Sugar and colorant adulteration	Rice syrup – 25	E150c – 50 ppm
	Rice syrup – 25	E150c – 100 ppm
	Rice syrup – 25	E150c – 150 ppm
	Rice syrup – 25	E150d – 50 ppm
	Rice syrup – 25	E150d – 100 ppm
	Rice syrup – 25	E150d – 150 ppm

UV–Vis spectra data acquisition

In order to dissolve any crystals and obtain homogeneous samples, honeys were heated at 40 °C for 1 h in a water bath before spectra acquisition. Authentic and adulterated honey samples were diluted with double-distilled water to a concentration level of 3% (w/w) and the UV–Vis spectra acquisition was carried out with a spectrophotometer (UV–1700, Shimadzu Corporation, Japan) in the wavelength range of 190–900 nm at 0.5 nm interval. Likewise, sugar syrups were diluted to a concentration level of 3% as honey samples, while the colorants were diluted in double-distilled water at 100 mg/Kg. For the spectroscopic measurements, diluted honey samples and adulterants were placed in quartz cells and distilled water was used as blank. All measurements were performed at room temperature. Five specimen replicates were used for each sample and the average spectrum of each honey sample was used for further analysis and construction of the chemometric models.

Data pre-processing and multivariate statistical analysis

Different pre-processing methods, such as Savitzky–Golay smoothing, 1st and 2nd derivatives (with a second order polynomial and 11-point window size) and standard normal variate (SNV), were evaluated for the elimination of systematic variations on the baseline. For all spectra processing, the SNV and 1st derivative transformations were eventually used, as they produced the best results. PCA was used in the original and pre-processed data in order to visualize and explore the similarities and differences among the samples. For classification purposes, PCA was followed by the Random Forest (RF) algorithm, Partial Least Squares – Discriminant Analysis (PLS-DA), and Data Driven-Soft Independent Modelling of Class Analogies (DD-SIMCA) for predicting whether or not a sample was adulterated.

Generally, DD-SIMCA is a supervised classification method which assigns new objects to the class when the degree of similarity obeys the decision criteria [37]. RF is a commonly used machine learning algorithm that combines the output of several decision trees to reach a single result. RF is a non-parametric method that can handle non-linearities. As an ensemble method, RF aggregates the predictions of all decision trees into the most popular results, thus achieving good results in various applications [43]. PLS-DA is a classification technique that maximizes the covariance between the spectral profiles and the classes to which the samples belong [44]. In contrast to DD-SIMCA, PLS-DA uses information from both target and non-target objects to optimize the model. Whilst DD-SIMCA uses information only from the target class to optimize the model. Previous studies on modelling approaches state that PLS-DA is not the best option to tackle food authentication issues [45, 46]. This is because, among other reasons, it is hard or even impossible to anticipate the adulteration scenarios of a specific food product, hampering the development of a robust discriminant analysis model that correctly classifies authentic food products without representative non-target samples in the training stage. However, in the present study PLS-DA was employed to explore the spectral ranges

that maximizes the differences between the adulterated honey samples (target class) and the authentic samples.

For the discrimination of authenticity of the honeys, both one-class classification methods, i.e., DD-SIMCA, and binary class classification methods, i.e., RF and PLS-DA, were used. The data set was split into two groups, the training set for building the models and the test set for validating the trained models. Two different strategies were used to split the dataset intro training and test set since we use a one-class classification method and two binary class classification methods. Specifically, for DD-SIMCA, 80% of the adulterated samples was used in the training set and the remaining 20% of the adulterated samples along with the authentic samples were placed in the test set. On the other hand, for RF and PLS-DA, 80% of the adulterated samples and 80% of the authentic samples were used in the training set, while the remaining 20% of both classes were placed in the test set. Therefore, in the case of DD-SIMCA, we aimed to test its ability to predict whether a sample is adulterated or not, while in the case of RF and PLS-DA, the samples are categorized in adulterated or authentic. The models were cross-validated using random blocks validation with 10 iterations. For RF and PLS-DA approaches, classes for both target (adulterated), and non-target (authentic honey samples) samples were considered. In contrast, the DD-SIMCA approach was used as an outlier detection method, by modelling only target (adulterated honey) class and excluding authentic honey samples that do not met the similarity criteria. The significance level was set at 5% ($\alpha = 0.05$) for all models. The evaluation was performed utilizing accuracy, sensitivity, and specificity as figures of merit for DD-SIMCA, and accuracy, precision, recall, and F1-score for RF and PLS-DA. The evaluation criteria were calculated according to the following equations.

$$Accuracy = \frac{(TP + TN)}{Total} = \frac{(TP + TN)}{(TP + FP + FN + TN)} \quad (1)$$

$$Precision = \frac{TP}{Predicted\ positive} = \frac{TP}{(TP + FP)} \quad (2)$$

$$Recall\ or\ Sensitivity = \frac{TP}{Actual\ positive} = \frac{TP}{(TP + FN)} \quad (3)$$

$$Specificity = \frac{TN}{Actual\ negative} = \frac{TN}{(FP + TN)} \quad (4)$$

$$F1 - score = 2 \times \frac{(Precision \times Recall)}{(Precision + Recall)} \quad (5)$$

where TP: True Positives, TN: True Negatives, FP: False Positives, and FN: False Negatives from the confusion matrix.

In the rest of the paper, all values for the criteria refer to those achieved in the test set. In the case of the DD-SIMCA models, we calculated the sensitivity on both the training and test set since the test set included only 20% of the true positive samples (as the rest of the true positive samples were included in the training set).

The experimental study was performed in Python using the scikit toolbox [47] of accessing the implementation of PCA, RF, and PLS-DA. The DD-SIMCA models were built using a MATLAB code available at https://github.com/yzontov/dd-simca.

Results and discussion

Analysis of UV–Vis spectra

Representative UV–Vis spectra of an authentic honey sample, adulterated samples of the same honey, and pure adulterants are depicted in Fig. 1. The UV–Vis spectra of the authentic and adulterated samples exhibit differences in the wavelength range of 250–350 nm (Fig. 1). According to Roshan et al. [37], the ultraviolet – visible region between 200 and 500 nm contains information about various components, such as phenolics, flavonoids, and conjugated systems that absorb in this spectral range. Generally, the range of 250–350 nm is related to the absorbance of sugars (primarily glucose and fructose), amino acids (mainly tryptophan), proteins and phenolic compounds [36, 38, 48].

For the determination of sugar adulteration, the addition of two different sugar types to honey, namely rice syrup and invert sugar (hydrolysed sucrose), was examined. In order

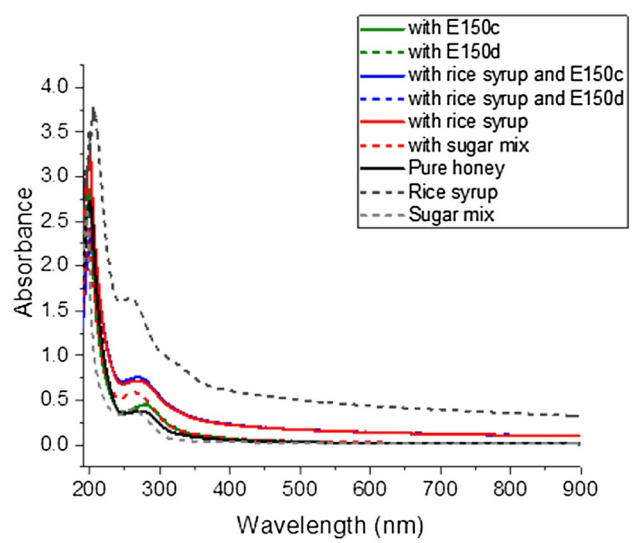

Fig. 1 The original UV–Vis spectra of authentic honey, pure adulterants, and adulterated samples with sugars and colorants in the highest concentration in the range of 190–900 nm

to investigate the difference between the types and quantify the percentage of adulteration, four Greek honey samples were selected, two light-coloured and two dark-coloured honeys. As illustrated in Fig. 2a, the UV–Vis spectrum of pure rice syrup results in higher absorbance than that of authentic honey samples, while the spectrum of pure invert sugar results in lower absorbance. The spectra of rice syrup and invert sugar demonstrated an absorbance band at around 260 nm. Therefore, as the level of adulteration increased, the peak height and the absorbance intensity also increased. The addition of syrups to the honey samples at low concentration (5 and 10% w/w) resulted in increased absorbance intensity at 280 nm, while the addition of higher concentration shifted the absorbance band in lower wavelengths in all tested honey samples (Fig. 2a). Specifically, the addition of rice syrup shifted the peak to 268 nm, while the addition of invert sugar to 265 nm. These results are in agreement with Guellis et al. [49], who studied polyfloral honey adulterated with corn syrup, and confirmed that the spectrum of syrup presented an intense band at 285 nm, whereas authentic honey showed a band with lower absorption intensity. On the other hand, Suhandy et al. [40], who studied the adulteration of Sumbawa honey with high fructose corn syrup, noticed that when syrups were added to the honeys at concentration level of 10%, the absorbance peak at 280 nm decreased. It is noted that these results are in agreement with the literature, as the absorbance band 200–250 nm is related to sugars, mainly glycose and fructose, present in honey [36].

For the determination of colorant adulteration, the addition of colorants, ammonia caramel (E150c) and sulfite ammonia caramel (E150d), was examined only in the case of light-coloured honey samples as this addition occurs only when the colour of the honey is very light. The two different colorants were added directly to the honey samples in three different concentrations, namely 50, 100, and 150 ppm. The addition of colorants resulted in a slight increase in the absorbance band at 279 nm. Small differences were observed in the spectra of the honeys adulterated with different levels of colorants; however, these differences do not seem to be considerable (Fig. 2b). The UV–Vis spectra of pure colorant E150d at a concentration level of 100 mg/Kg (diluted in water) displayed absorbance peaks at 220 and 276 nm, however these peaks are not clearly seen in the spectra of adulterated honey samples.

Lastly, the simultaneous addition of both rice syrup and colorants was examined. The two different colorants (E150c and E150d) were added into the rice syrup in three different concentrations, namely 50, 100, and 150 ppm, and the mixture was added to the honey samples at a ratio of 25%. As noted above, the addition of rice syrup (at a ratio of 25%) increased the absorbance and shifted the band slightly to lower wavelengths, i.e., at 270 nm for all samples. In most cases, adding a different amount of colorant did not cause significant peak shifting and changes in absorbance intensity (Fig. 2c). Hence, we can conclude that the addition of syrup resulted in UV–Vis spectra with higher intensities than those with the addition of colorants.

Principal component analysis

PCA was applied to differentiate the honeys into authentic and adulterated samples. For this reason, the range of 220–550 nm was selected for further analysis, as this range contains the most useful information [50]. As depicted in Fig. 3a there was no trend of clear separation between the adulterated and the authentic thyme honey samples using original UV–Vis spectra, except for those adulterated with higher concentration of sugars. Specifically, according to the scatter plot of the first two PCs, the adulterated samples with the simultaneous addition of rice syrup and colorants (red circles) are clearly separated from the authentic thyme samples which are depicted with yellow circles (Fig. 3a). On the other hand, the samples adulterated only with syrups (blue circles), depending on the concentration of the adulteration,

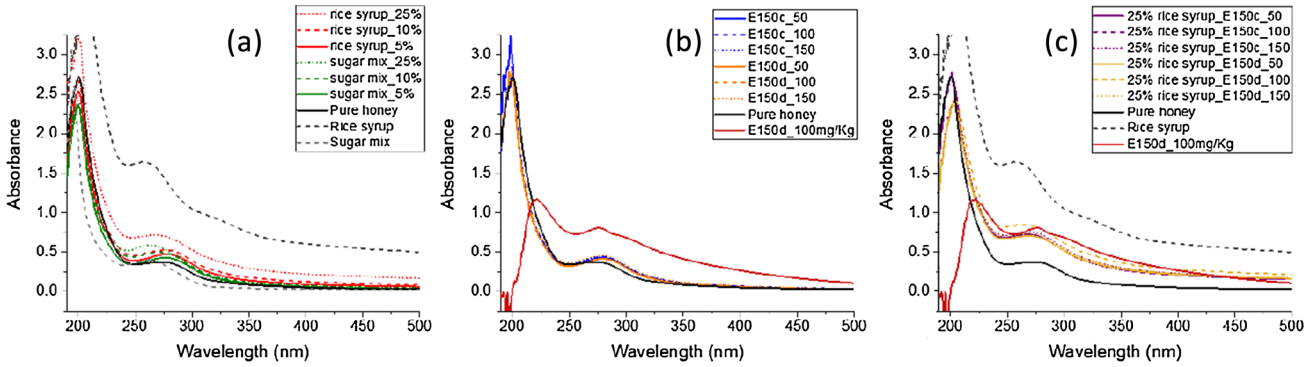

Fig. 2 The original UV–Vis spectra of authentic honey, pure adulterants, and adulterated samples with (**a**) syrups, (**b**) colorants, and (**c**) both syrups and colorants in different concentrations in the range of 190–500 nm

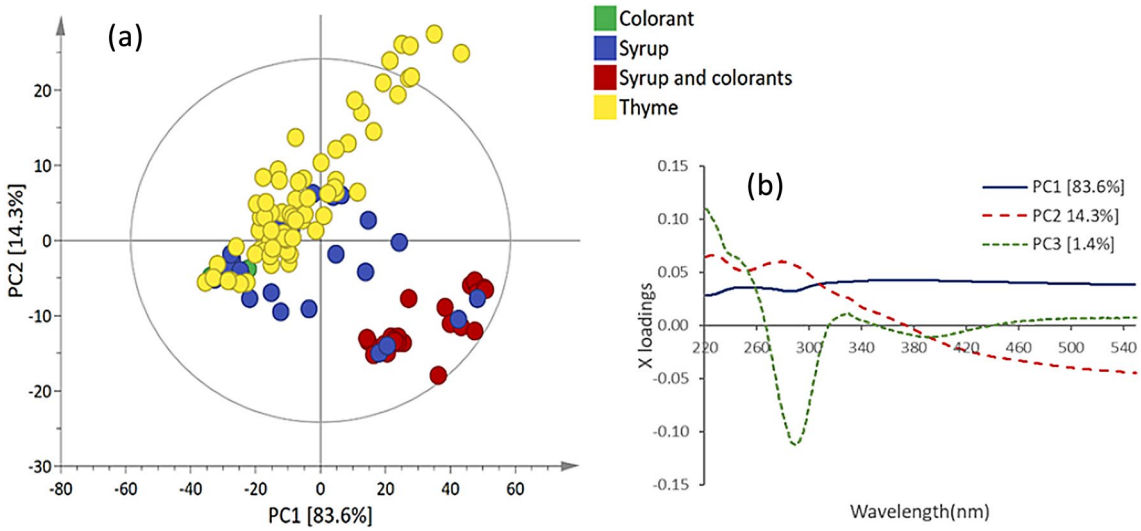

Fig. 3 PCA score plots (**a**) and loading plots of first three components (**b**) of the authentic and adulterated thyme honey samples using the original UV–Vis spectra in the range of 220–550 nm (yellow circles: authentic thyme honey samples, blue circles: adulterated samples with syrups, green circles: adulterated samples with colorants, red circles: adulterated samples with both syrup and colorants)

are intermediate between the two previous groups, while the samples to which colorants were added (green circles) are located in the area of authentic honeys. According to the PCA, the total explained variance for the first three PCs was 99.3%, with the first and second PCs accounted for 86.3 and 14.3%, respectively. The loading plot (Fig. 3b) indicates that several spectrum bands at around 240, 250, 278, 285, 290, and 330 nm seem to be the most influential variables to classify honey samples. In order to remove baseline shifts for obtaining better results, pre-processing techniques, SNV and 1st derivative were applied for models' construction.

Results from predictive models

Given that the adulteration is defined as the addition of any adulterant to honey, regardless of its type and concentration [2], RF, PLS-DA, and DD-SIMCA models were constructed in order to identify if a honey is adulterated or authentic using (i) only thyme honeys and (ii) all honey samples in the validation set. Considering all types of adulterated samples in the training set, RF, PLS-DA, and DD-SIMCA models achieved an accuracy of 92, 92, and 89%, respectively, when only thyme honeys were used in the validation set. The aforementioned models achieved 89, 92, and 97% of sensitivity/recall in the validation set, which reflects the number of correct positive decisions divided by the total number of positive cases, and therefore these models can identify the adulterated samples with high accuracy. More specifically, in the case of DD-SIMCA models, only one adulterated sample with low concentration of colorant addition was misclassified, while eleven authentic honeys were misclassified as adulterated (Fig. 4(1b)). The predictive model misclassified nine Greek thyme samples out of 26, while only one out of 31 from Tunisia and one out of 3 from Turkey. These results were expected because only Greek honeys were used for adulteration and, hence, the model recognized similarities in the spectrum with Greek authentic honeys. In our previous work, which aimed to distinguish the geographical and botanical origin of honeys, a good differentiation of thyme honeys among samples from different geographical origins was performed [50]. Thus, the prediction of the present models could be increased if thyme honeys from other countries are also used in the adulterated honey data set. It is worth noting that it is more important in most applications for a model to correctly identify the majority of adulterated samples even if it misclassifies a proportion of authentic ones (i.e., aiming for larger recall/sensitivity than specificity in the case of DD-SIMCA), than the opposite. This may add some extra effort in cases that the model falsely misclassifies unadulterated honey samples, as additional tests need to be performed to detect the authenticity of the honeys by food authorities, but it will avoid making wrong estimates about adulterated samples.

Based on the aforementioned good performance of these models, polyfloral and pine honeys were also added in the data set. RF, PLS-DA, and DD-SIMCA models achieved an accuracy of 95, 88, and 93%, respectively. RF and DD-SIMCA resulted in an increased accuracy when utilizing a larger data set (Table 2). As illustrated in the acceptance plot of the validation set in Fig. 4(1c), the DD-SIMCA model misclassified only one adulterated sample out of 58, while seventeen authentic samples out of 151 were predicted as adulterated, of which fourteen were from Greece (ten thyme and four polyfloral honeys). These results are in agreement

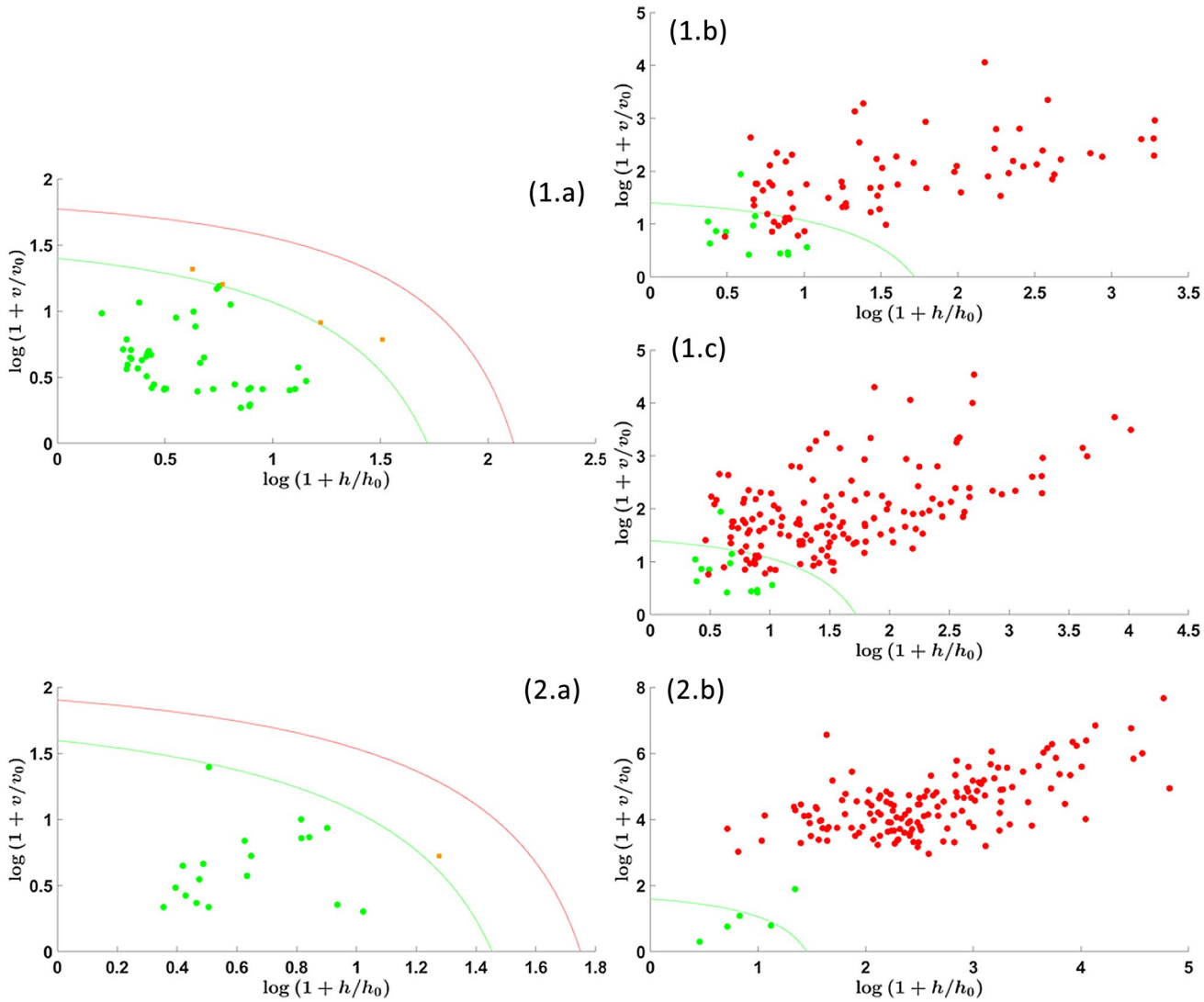

Fig. 4 Acceptance plot of the training set (1**a** and 2**a**,) and validation set (1**b**, 1**c**, and 2**b**) obtained by DD-SIMCA models using UV–Vis spectra, pre-processed with SNV and 1st derivative, for detection of adulteration using only thyme samples (1**a** and 1**b**) and all honey samples (1**a** and 1**c**) in the validation set and adulteration with simultaneous addition of syrup and colorants (2**a** and 2**b**) using all honey samples in the validation set. [Acceptance plot for training set: training samples are illustrated in green circles, while extreme objects are illustrated in orange circles between the green and red lines. Acceptance plot for validation set: the honey samples for target class are illustrated in green circles, while samples for non-target class are depicted in red circles.]

with de Souza et al. [39] who applied DD-SIMCA models, using a wavelength range of 320–800 nm, and confirmed that the adulteration of honey with corn syrup, agave syrup, and sugarcane molasses could be identified, achieving a 100% of sensitivity and specificity. Moreover, Suhandy *et al*. [40], who used only PCA, observed a visual separation between authentic and adulterated Sumbawa honeys with high fructose corn syrup, using a wavelength range of 230–400 nm.

Regarding the identification of the type of the honey adulteration, honeys adulterated with both syrup and colorants were used to build the models, while the authentic honeys were used for validation. RF, PLS-DA, and DD-SIMCA models achieved an accuracy of 98, 99, and 99% using only thyme honey samples in the validation set, respectively, while an accuracy of 97, 99, and 93% was attained using all honey samples for validation. According to the results from DD-SIMCA models, only one adulterated honey sample (which contained the colorant E150d at 50 ppm) was misclassified, while all authentic honeys were correctly categorized as unadulterated, as depicted in Fig. 4(2b).

With regards to sugar adulteration, the results from the developed models showed a good performance, achieving an accuracy of 91, 92, and 94% for RF, PLS-DA, and DD-SIMCA, respectively, using all honey samples in the

Table 2 The best results obtained from Random Forest, PLS-DA, and DD-SIMCA models for identification of honey adulteration using UV–Vis spectra

	RF				PLS-DA				DD-SIMCA		
	Accuracy %	Precision %	Recall %	F1-score %	Accuracy %	Precision %	Recall %	F1-score %	Accuracy %	Sensitivity %	Specificity %
Identification of adulterated honey samples – using only thyme honeys in the validation set											
Adulterated or not	91.9	93.6	89.2	91.0	91.5	92.0	90.8	90.9	89.3	97.4	89.6
Adulterated with both syrup and colorants	98.4	100	94.0	96.7	99.0	96.7	100	98.2	99.0	96.7	100
Adulterated with syrup	88.4	81.0	74.0	76.3	92.1	90.7	80.0	83.9	89.1	90.9	91.3
Adulterated with syrup (either only syrup or syrup and colorant)	88.7	89.3	82.2	84.6	91.3	87.6	92.2	89.4	93.8	93.6	97.2
Adulterated with colorant	97.5	100	80.0	86.7	95.0	78.3	85.0	81.0	98.4	89.1	100
Adulterated with colorant (either only colorant or colorant and syrup)	97.6	97.3	95.7	96.4	97.1	93.8	98.6	95.9	97.7	95.1	99.9
Identification of adulterated honey samples – using all honeys in the validation set											
Adulterated or not	94.8	94.3	98.7	96.4	88.3	87.6	98.0	92.4	93.3	97.4	93.8
Adulterated with both syrup and colorants	97.4	98.0	84.0	89.7	99.7	98.3	100	99.1	99.5	96.7	100
Adulterated with syrup	91.1	80.5	56.0	62.8	92.0	80.7	52.0	60.4	93.8	90.9	95.0
Adulterated with syrup (either only syrup or syrup and colorant)	91.3	85.9	75.6	79.8	91.8	91.8	71.1	79.7	96.2	93.6	98.0
Adulterated with colorant	97.8	86.7	90.0	84.7	91.6	20.0	20.0	18.7	99.2	89.1	100
Adulterated with colorant (either only colorant or colorant and syrup)	98.4	100	91.4	95.3	94.9	95.0	77.1	83.9	98.9	95.1	99.9

validation set. According to the results, some adulterated samples with invert sugar were misclassified as authentic ones, possibly, due to the fact this added sugar is similar in carbohydrates composition than honey (glucose and fructose). On the other hand, rice syrup has a different composition with maltose and maltotriose as main carbohydrates. When all samples that were adulterated with syrups (all honey samples that were adulterated either with syrup or syrup and colorant) were used in the training set, the accuracy of the models increased for all cases considered (Table 2). This is reasonable because more samples, with higher concentration of syrup were added for training the models, and thus, the sensitivity and specificity of the models were also increased.

In relation to the colorant adulteration, two types of colorants at three concentration levels were considered. RF, PLS-DA, and DD-SIMCA achieved an accuracy of 98, 92, and 99%, respectively, when considering all honey samples in the validation set. Despite its good accuracy, PLS-DA fails to recognize the adulterated samples, and this leads to poor performance in terms of the other evaluation criteria (precision, recall, and F1-score are approximately equal to 20%). Similar results (98, 95, and 98% for RF, PLS-DA, and DD-SIMCA, respectively) were obtained for the case where only the thyme honey samples were included in the validation set. When all samples that were adulterated with colorants (all honey samples that were adulterated either with only colorant or colorant and syrup) were used in the training set, the accuracy of the models slightly increased in most cases, as more samples were added into the training set.

All predictive models achieved good results (in the majority of cases an accuracy of more than 90% was achieved). DD-SIMCA was the winner in terms of accuracy in four out of six cases when all honey samples were included in the validation set and three out of six cases when only thyme honey samples were included in the validation set. Similar conclusions were drawn if the comparison criterion is recall/sensitivity.

Conclusions

In the present work, UV–Vis spectroscopy combined with multivariate statistical analysis was used in order to identify honey adulteration with sugar syrups and colorants. The results showed that UV–Vis spectra processed with chemometrics were successfully used to identify the adulteration of thyme and other Mediterranean honeys, with high accuracy. The wavelength range of 220–550 nm was utilized for the models' construction, since PCA of the spectra showed that the wavelengths at around 240, 250, 278, 285, 290, and 330 nm are the most useful for discrimination of honey adulteration. RF, PLS-DA, and DD-SIMCA models were developed to identify honey adulteration and the prediction accuracy was over 90% for most models. DD-SIMCA achieved better accuracy and recall/sensitivity scores than RF and PLS-DA in most cases. Since comparing one class vs binary class classification methods is not straightforward, we conclude that DD-SIMCA is more suitable in the case of the application needs to maximize the sensitivity, while RF is more suitable for applications that prioritize the specificity. Hence, the results from the present study support that UV–Vis spectroscopy has great potential for identifying honey adulteration in a quick and non-destructive manner, and also having the advantage of being a simple, inexpensive, and fast method. However, further studies with a larger number of honey samples are necessary, in order to develop and verify robust models.

Acknowledgements This paper is part of the PRIMA programme, supported by European Union's Horizon 2020 research and innovation programme, under grant agreement No 1932, project MEDIFIT (Call 2019 Section "Introduction" Agrofood IA).

Author contributions Investigation: DDP; Formal analysis: DDP, NP, EK; Methodology: DDP, CSS, SS, EK; Resources: DDP, CSS, SS, EK, KK, VV; Writing—original draft preparation: DDP; Writing—review and editing: EK, NP, KK, SS, CSS, VV; Funding acquisition: KK; Project administration: EK, KK, VV; Supervision: EK.

Funding This study was supported by Horizon Europe, No 1932.

Data availability The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declare no conflict of interest.

Compliance with ethics requirements This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Samarghandian S, Farkhondeh T, Samini F (2017) Honey and health: a review of recent clinical research. Pharmacogn Res 9(2):121. https://doi.org/10.4103/0974-8490.204647
2. EU Council Directive 2001/110/EC of 20 December 2001 relating to honey. Off J Eur Commun Legis 10:47–52 (2002) Available: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32001L0110&from=EN
3. Machado DeMelo AA, AlmeidaMuradian LBd, Sancho MT, PascualMaté A (2018) Composition and properties of *Apis mellifera* honey: a review. J Apic Res 57(1):5–37. https://doi.org/10.1080/00218839.2017.1338444
4. Pita-Calvo C, Vázquez M (2017) Differences between honeydew and blossom honeys: a review. Trends in Food Sci Technol 59:79–87. https://doi.org/10.1016/j.tifs.2016.11.015
5. Cengiz MF, Durak MZ (2019) Rapid detection of sucrose adulteration in honey using Fourier transform infrared spectroscopy.

Spectrosc Lett 52(5):267–273. https://doi.org/10.1080/00387010.2019.1615957
6. Fakhlaei R, Selamat J, Khatib A, Razis AFA, Sukor R, Ahmad S, Babadi AA (2020) The toxic impact of honey adulteration: a review. Foods 9(11):1538
7. Wang J, Kliks MM, Jun S, Jackson M, Li QX (2010) Rapid analysis of glucose, fructose, sucrose, and maltose in honeys from different geographic regions using Fourier transform infrared spectroscopy and multivariate analysis. J Food Sci 75(2):C208–C214. https://doi.org/10.1111/j.1750-3841.2009.01504.x
8. Sivakesava S, Irudayaraj J (2001) Prediction of inverted cane sugar adulteration of honey by Fourier transform infrared spectroscopy. J Food Sci 66(7):972–978. https://doi.org/10.1111/j.1365-2621.2001.tb08221.x
9. Tewari J, Irudayaraj J (2004) Quantification of saccharides in multiple floral honeys using Fourier transform infrared microattenuated total reflectance spectroscopy. J Agric Food Chem 52(11):3237–3243. https://doi.org/10.1021/jf035176+
10. Kelly JD, Downey G, Fouratier V (2004) Initial study of honey adulteration by sugar solutions using midinfrared (MIR) spectroscopy and chemometrics. J Agric Food Chem 52(1):33–39. https://doi.org/10.1021/jf034985q
11. Siddiqui AJ, Musharraf SG, Choudhary MI, Rahman A-U (2017) Application of analytical methods in authentication and adulteration of honey. Food Chem 217:687–698. https://doi.org/10.1016/j.foodchem.2016.09.001
12. Das C, Chakraborty S, Acharya K, Bera NK, Chattopadhyay D, Karmakar A, Chattopadhyay S (2017) FT-MIR supported electrical impedance spectroscopy based study of sugar adulterated honeys from different floral origin. Talanta 171:327–334. https://doi.org/10.1016/j.talanta.2017.05.016
13. Huang F, Song H, Guo L, Guang P, Yang X, Li L, Zhao H, Yang M (2020) Detection of adulteration in Chinese honey using NIR and ATR-FTIR spectral data fusion. Spectrochim Acta A Mol Biomol Spectrosc 235:118297. https://doi.org/10.1016/j.saa.2020.118297
14. Riswahyuli Y, Rohman A, Setyabudi FMCS, Raharjo S (2020) Indonesian wild honey authenticity analysis using attenuated total reflectance-fourier transform infrared (ATR-FTIR) spectroscopy combined with multivariate statistical techniques. Heliyon 6(4):e03662. https://doi.org/10.1016/j.heliyon.2020.e03662
15. Se KW, Wahab RA, Syed Yaacob SN, Ghoshal SK (2019) Detection techniques for adulterants in honey: challenges and recent trends. J Food Compos Anal 80:16–32. https://doi.org/10.1016/j.jfca.2019.04.001
16. Beretta G, Granata P, Ferrero M, Orioli M, Maffei Facino R (2005) Standardization of antioxidant properties of honey by a combination of spectrophotometric/fluorimetric assays and chemometrics. Anal Chim Acta 533(2):185–191. https://doi.org/10.1016/j.aca.2004.11.010
17. Gan Z, Yang Y, Li J, Wen X, Zhu M, Jiang Y, Ni Y (2016) Using sensor and spectral analysis to classify botanical origin and determine adulteration of raw honey. J Food Eng 178:151–158. https://doi.org/10.1016/j.jfoodeng.2016.01.016
18. Wang S, Guo Q, Wang L, Lin L, Shi H, Cao H, Cao B (2015) Detection of honey adulteration with starch syrup by high performance liquid chromatography. Food Chem 172:669–674. https://doi.org/10.1016/j.foodchem.2014.09.044
19. Yilmaz MT, Tatlisu NB, Toker OS, Karaman S, Dertli E, Sagdic O, Arici M (2014) Steady, dynamic and creep rheological analysis as a novel approach to detect honey adulteration by fructose and saccharose syrups: correlations with HPLC-RID results. Food Res Int 64:634–646. https://doi.org/10.1016/j.foodres.2014.07.009
20. Ruiz-Matute AI, Soria AC, Martínez-Castro I, Sanz ML (2007) A new methodology based on GC−MS to detect honey adulteration with commercial syrups. J Agric Food Chem 55(18):7264–7269. https://doi.org/10.1021/jf070559j
21. Bertelli D, Lolli M, Papotti G, Bortolotti L, Serra G, Plessi M (2010) Detection of honey adulteration by sugar syrups using one-dimensional and two-dimensional high-resolution nuclear magnetic resonance. J Agric Food Chem 58(15):8495–8501. https://doi.org/10.1021/jf101460t
22. Rachineni K, Rao Kakita VM, Awasthi NP, Shirke VS, Hosur RV, Chandra Shukla S (2022) Identifying type of sugar adulterants in honey: combined application of NMR spectroscopy and supervised machine learning classification. Curr Res Nutr Food Sci 5:272–277. https://doi.org/10.1016/j.crfs.2022.01.008
23. Padovan GJ, De Jong D, Rodrigues L, Marchini JS (2003) Detection of adulteration of commercial honey samples by the 13C/12C isotopic ratio. Food Chem 82(4):633–636. https://doi.org/10.1016/S0308-8146(02)00504-6
24. Tosun M (2013) Detection of adulteration in honey samples added various sugar syrups with 13C/12C isotope ratio analysis method. Food Chem 138(2):1629–1632. https://doi.org/10.1016/j.foodchem.2012.11.068
25. Başar B, Özdemir D (2018) Determination of honey adulteration with beet sugar and corn syrup using infrared spectroscopy and genetic-algorithm-based multivariate calibration. J Sci Food Agric 98(15):5616–5624. https://doi.org/10.1002/jsfa.9105
26. Li Q, Zeng J, Lin L, Zhang J, Zhu J, Yao L, Wang S, Yao Z, Wu Z (2020) Low risk of category misdiagnosis of rice syrup adulteration in three botanical origin honey by ATR-FTIR and general model. Food Chem 332:127356. https://doi.org/10.1016/j.foodchem.2020.127356
27. Bázár G, Romvári R, Szabó A, Somogyi T, Éles V, Tsenkova R (2016) NIR detection of honey adulteration reveals differences in water spectral pattern. Food Chem 194:873–880. https://doi.org/10.1016/j.foodchem.2015.08.092
28. Guelpa A, Marini F, du Plessis A, Slabbert R, Manley M (2017) Verification of authenticity and fraud detection in South African honey using NIR spectroscopy. Food Control 73:1388–1396. https://doi.org/10.1016/j.foodcont.2016.11.002
29. Oroian M, Ropciuc S, Paduret S (2018) Honey adulteration detection using Raman spectroscopy. Food Anal Methods 11(4):959–968. https://doi.org/10.1007/s12161-017-1072-2
30. Rios-Corripio MA, Rios-Leal E, Rojas-López M, Delgado-Macuil R (2011) FTIR characterization of Mexican honey and its adulteration with sugar syrups by using chemometric methods. J Phys Conf Ser 274:012098. https://doi.org/10.1088/1742-6596/274/1/012098
31. Wu X, Xu B, Ma R, Niu Y, Gao S, Liu H, Zhang Y (2022) Identification and quantification of adulterated honey by Raman spectroscopy combined with convolutional neural network and chemometrics. Spectrochim Acta A Mol Biomol Spectrosc 274:121133. https://doi.org/10.1016/j.saa.2022.121133
32. Chen Q, Qi S, Li H, Han X, Ouyang Q, Zhao J (2014) Determination of rice syrup adulterant concentration in honey using three-dimensional fluorescence spectra and multivariate calibrations. Spectrochim Acta A Mol Biomol Spectrosc 131:177–182. https://doi.org/10.1016/j.saa.2014.04.071
33. Yan S, Sun M, Wang X, Shan J, Xue X (2022) A novel, rapid screening technique for sugar syrup adulteration in honey using fluorescence spectroscopy. Foods 11(15):2316. https://doi.org/10.3390/foods11152316
34. Ansari MJ, Al-Ghamdi A, Khan KA, Adgaba N, El-Ahmady SH, Gad HA, Roshan A, Meo SA, Kolyali S (2018) Validation of botanical origins and geographical sources of some Saudi honeys using ultraviolet spectroscopy and chemometric analysis. Saudi J Biol Sci 25(2):377–382. https://doi.org/10.1016/j.sjbs.2017.09.005

35. Nunes A, Zilto Azevedo G, dos Rocha Santos B, Vanz Borges C, Pace Pereira Lima G, Conte Crocoli L, Moura S, Maraschin M (2022) Characterization of Brazilian floral honey produced in the states of Santa Catarina and São Paulo through UltraViolet–Visible (UV–Vis), near-infrared (NIR), and nuclear magnetic resonance (NMR) spectroscopy. Food Res Int 162:111913. https://doi.org/10.1016/j.foodres.2022.111913
36. Orfanakis E, Markoulidakis M, Philippidis A, Zoumi A, Velegrakis M (2021) Optical spectroscopy methods combined with multivariate statistical analysis for the classification of cretan thyme, multi-floral and honeydew honey. J Sci Food Agric 101(13):5337–5347. https://doi.org/10.1002/jsfa.11182
37. Roshan A-RA, Gad HA, El-Ahmady SH, Khanbash MS, Abou-Shoer MI, Al-Azizi MM (2013) Authentication of monofloral Yemeni sidr honey using ultraviolet spectroscopy and chemometric analysis. J Agric Food Chem 61(32):7722–7729. https://doi.org/10.1021/jf402280y
38. Suhandy D, Yulia M (2021) The use of UV spectroscopy and SIMCA for the authentication of indonesian honeys according to botanical, entomological and geographical origins. Molecules 26(4):915. https://doi.org/10.3390/molecules26040915
39. de Souza RR, de Sousa Fernandes DD, Diniz PHGD (2021) Honey authentication in terms of its adulteration with sugar syrups using UV–Vis spectroscopy and one-class classifiers. Food Chem 365:130467. https://doi.org/10.1016/j.foodchem.2021.130467
40. Suhandy D, Kusumiyati Kuncoro S, Rahmawati W, Yulia M (2022) Authentication of pure and adulterated Sumbawa monofloral honey using ultraviolet-visible spectroscopy. IOP Conference Series: Earth and Environmental Science, 1027(1): 012014. doi: https://doi.org/10.1088/1755-1315/1027/1/012014
41. Mitra PK, Karmakar R, Nandi R, Gupta S (2023) Low-cost rapid workflow for honey adulteration detection by UV–Vis spectroscopy in combination with factorial design, response surface methodology and supervised machine learning classifiers. Bioresour Technol Rep 21:101327. https://doi.org/10.1016/j.biteb.2022.101327
42. Valinger D, Longin L, Grbeš F, Benković M, Jurina T, Kljusurić JG, Tušek AJ (2021) Detection of honey adulteration–the potential of UV-Vis and NIR spectroscopy coupled with multivariate analysis. LWT 145:111316. https://doi.org/10.1016/j.lwt.2021.111316
43. Breiman L (2001) Random forests. Mach Learn 45(1):5–32. https://doi.org/10.1023/A:1010933404324
44. Barker M, Rayens W (2003) Partial least squares for discrimination. J Chemom A J Chemom Soc 17(3):166–173. https://doi.org/10.1002/cem.785
45. Rodionova OY, Titova AV, Pomerantsev AL (2016) Discriminant analysis is an inappropriate method of authentication. TrAC Trends Anal Chem 78:17–22. https://doi.org/10.1016/j.trac.2016.01.010
46. Rodionova OY, Pomerantsev AL (2020) Chemometric tools for food fraud detection: the role of target class in non-targeted analysis. Food chem 317:126448. https://doi.org/10.1016/j.foodchem.2020.126448
47. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M et al (2011) Scikit-learn: machine learning in Python. J Mach Learn Res 12:2825–2830. https://doi.org/10.48550/arXiv.1201.0490
48. Parri E, Santinami G, Domenici V (2020) Front-face fluorescence of honey of different botanic origin: a case study from Tuscany (Italy). Appl Sci 10(5):1776. https://doi.org/10.3390/app10051776
49. Guellis C, Valério DC, Bessegato GG, Boroski M, Dragunski JC, Lindino CA (2020) Non-targeted method to detect honey adulteration: combination of electrochemical and spectrophotometric responses with principal component analysis. J Food Compos Anal 89:103466. https://doi.org/10.1016/j.jfca.2020.103466
50. Dimakopoulou-Papazoglou D, Ploskas N, Koutsoumanis K, Katsanidis E. Identification of geographical and botanical origin of Mediterranean honeys using UV-vis spectroscopy and multivariate statistical analysis. (Under review)