

1 **A SEMI-QUANTITATIVE ONE-CLASS PARTIAL LEAST SQUARES MODEL FOR DETECTING HONEY**
2 **ADULTERATION USING TD-NMR SPECTROSCOPY**

3 **Authors' names**

4 Glòria Rovira^a, Carolina Sheng Whei Miaw Botelho^b, Laura Lima de Oliveira ^c, Marcus Vinicius de Oliveira
5 Andrade ^d, Poliana Macedo Santos^e, Marcelo Martins Sena ^f, Scheilla Vitorino Carvalho de Souza ^b, Maria
6 Pilar Callao^a, Itziar Ruisánchez ^{a,*}

7 ^a – Chemometrics and Sensorics for Analytical Solutions (CHEMOSENS) Group, Department of Analytical
8 and Organic Chemistry, Rovira i Virgili University, Marcel·lí Domingo s/n, 43007, Tarragona, Spain.

9 ^b - Food Science Graduate Program (PPGCA), Faculty of Pharmacy (FAFAR), Federal University of Minas
10 Gerais (UFMG), Av. Antônio Carlos, 6627, Campus da UFMG, Pampulha, 31270-010, Belo Horizonte, MG,
11 Brazil.

12 ^c - Faculty of Pharmacy (FAFAR), Federal University of Minas Gerais (UFMG), Av. Antônio Carlos, 6627,
13 Campus da UFMG, Pampulha, 31270-010, Belo Horizonte, MG, Brazil.

14 ^d- Brazilian Federal Police, Regional Superintendence in Minas Gerais, 30441-170, Belo Horizonte, MG,
15 Brazi.

16 ^e - Chemistry and Biology Department (DAQBI), Federal Technological University of Paraná (UTFPR), Rua
17 Dep. Heitor Alencar Furtado, 5000, Cidade Industrial, 81280-340, Curitiba, PR, Brazil.

18 ^f - Chemistry Department, Institute of Exact Sciences (ICEx), Federal University of Minas Gerais (UFMG),
19 Av. Antônio Carlos, 6627, Campus da UFMG, Pampulha, 31270-010, Belo Horizonte, MG, Brazil. Instituto
20 Nacional de Ciência e Tecnologia em Bioanalítica (INCT-Bio), Campinas, SP, 13083-970, Brazil.

21

22 **Keywords:**

23 Multivariate classification, Time-Domain NMR, Honey adulteration, Semi-quantitative
24 performance parameters, Uncertainty interval, Decision limit

25

26

27

28

Preprint not peer reviewed

29 **Abstract**

30 This study developed a screening method to detect honey adulteration with inverted sugar
31 syrup using two types of honey samples; eucalyptus (EUC) and wild (W). For each type, 35
32 authentic and 42 adulterated samples were analyzed using Time-Domain Nuclear Magnetic
33 Resonance (TD-NMR). One-class Partial Least Squares Discriminant Analysis (OCPLS) models
34 were built from the authentic samples of each honey type, with performance evaluated via
35 sensitivity and specificity. Semi-quantitative performance parameters were estimated by
36 Performance Characteristic Curves (PCC). Decision limits ($CC\alpha$) were set at 5% (EUC) and 2% (W)
37 adulteration, below which adulterated samples could not be distinguished from authentic ones.
38 Two class limits were defined for unknown sample predictions, creating an uncertainty region
39 (UR). Samples within the UR were assigned as inconclusive and should be submitted for further
40 analysis. Among non-adulterated samples, 17% (EUC) and 43% (W) were classified as
41 inconclusive, avoiding false adulteration assignments. For adulterated samples, 19% (EUC) and
42 7% (W) fell in the UR, avoiding false non-adulteration assignments. This approach eliminated
43 classification errors, ensuring 100% reliability in identifying both non-adulterated and
44 adulterated samples. The method offers a robust tool for accurate honey authentication,
45 obtaining inconclusive samples that require further confirmation.

46

47 **1. Introduction**

48

49 According to the Standard for Honey CXS 12-1981 from the Codex Alimentarius, honey is defined
50 as the natural sweet substance produced by honeybees from the nectar of plants or secretions
51 of living parts of plants (Codex Alimentarius, 2001). Honey is widely recognized for its good and
52 attractive organoleptic properties and high health benefits (Se et al., 2019; Huang et al., 2020).

53 The price of high-quality honey has gone up due to the rising demand and limited availability.
54 Consequently, honey is susceptible to adulteration (Ciursă et al., 2021). This type of fraud
55 violates the standards established by Codex Alimentarius (Codex Alimentarius, 2001) and
56 regulatory bodies such as the European Commission (De Souza et al., 2021; European
57 Commission, 2018) and the Brazilian Ministry of Agriculture and Livestock (Instrução Normativa
58 Nº 11, 2000), due to the possible adverse health impact of the consumption of adulterated
59 honey (Fakhlaei et al., 2020).

60 There are different types of fraud in honey but the most common is the direct addition of syrups.
61 The most used syrups are high-fructose corn syrup (Başar et al., 2018; Li et al., 2017), rice syrup
62 (Li et al., 2020; Limm et al., 2023), inverted sugar (Ciursă et al., 2021), cane sugar (De Souza et
63 al., 2021; Dias et al., 2022), and other sugar syrups (Huang et al., 2020; Brar et al., 2023; Egidio
64 et al., 2024).

65 In this work, a strategy was established to detect the adulteration of honey by the addition of
66 inverted sugar syrup, one of the most common adulterants reported in fraud (Ciursă et al.,
67 2021). The official reference method for the identification of syrup-adulterated honey is based
68 on stable carbon isotope ratio analysis (AOAC, 1998). This method is time-consuming, sample-
69 destructive, and requires expensive instruments. In recent years, an effort has been made to
70 implement faster, environmentally friendly, and non-destructive techniques, such as molecular
71 spectroscopy (Huang et al., 2020). To cite some examples of spectroscopic techniques combined
72 with multivariate analysis in honey adulteration: near-infrared (NIR) (Huang et al., 2020; Fakhlaei
73 et al., 2020; Guelpa et al., 2017; Biaswas et al., 2024), UV-visible (De Souza et al., 2021), Fourier-
74 transform infrared-attenuated total reflectance (FTIR-ATR) (Ciursă et al., 2021; Cárdenas-
75 Escudero et al., 2023; Sahland et al., 2019), spectrofluorimetry (Antônio et al., 2022), Raman
76 (Oroian et al., 2018), and nuclear magnetic resonance (NMR) (Rachineni et al., 2022; Ribeiro et
77 al., 2014; Biaswas et al., 2023). In recent years, a type of NMR technique named Time-domain

78 nuclear magnetic resonance (TD-NMR) has been investigated in food science. In TD-NMR studies
79 proton relaxation is characterized by the longitudinal relaxation time (T_1) and transverse
80 relaxation time (T_2). T_2 relaxation time is multiexponential and indicates the presence of water
81 populations in food matrices (Ribeiro et al., 2014; Bertram et al., 2001; Belton, 1990; Finch et
82 al., 1971). TD-NMR offers different information about ^1H nuclei such as characterizing bound
83 water, free water, and their exchange between these states (Santos et al., 2016). This technique
84 requires minimal sample preparation, presents simple and fast measurement procedures, and
85 is non-invasive, which makes it suitable for application in food control. Some examples of its
86 application include detecting food adulteration in products such as cheese (Curti et al., 2023;
87 Machado et al., 2022), milk (Santos et al., 2016; Coimbra et al., 2020), honey (Ribeiro et al.,
88 2014), fruits (Colnago et al., 2021), and oil (Cengiz, 2023; Ribeiro et al., 2021).

89 Qualitative methods relying on multivariate classification techniques, such as partial least
90 squares discriminant analysis (PLS-DA) (Antônio et al., 2022; Brereton et al., 2014), soft
91 independent modelling of class analogy (SIMCA) (De Souza et al., 2021), and one-class partial
92 least squares discriminant analysis (OCPLS) (De Souza et al., 2021), have increasingly been
93 employed for detecting food fraud. The validation of these classification methods typically
94 involves assessing main performance parameters, such as sensitivity, specificity, accuracy,
95 precision, and occurrence (López et al., 2015; Cuadros-Rodríguez et al. 2016; Jiménez-Carvelo et
96 al., 2020). These parameters are calculated from the four binary possible responses: True
97 Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) indicating whether
98 a sample belongs to or does not belong to the modelled class. To determine if a sample belongs
99 to the modelled class, it is necessary to set the limit for this class, which is usually defined as a
100 model distance. Samples with a distance value lower than the class limit are considered to fit
101 within the class model (Jiménez-Carvelo et al., 2020; Rovira et al., 2023; Ruisánchez et al., 2021).

102 The performance parameters mentioned above refer to a binary response, e.g., the unknown
103 sample being compliant or not. When dealing with adulteration problems, it is important to
104 consider that the model's ability will depend on the level of sample adulteration. Therefore, in
105 addition to strictly qualitative performance parameters, concentration-related performance
106 (semi-quantitative) parameters should be established. Recently, the detection limit ($CC\alpha$), the
107 detection capability ($CC\beta$), and the unreliability region (UR) have been referenced (Ruisánchez
108 et al., 2022) with their estimates derived from performance characteristic curves (PCC)
109 (Ruisánchez et al., 2022; Ricardo et al., 2021; Corps et al., 2021).

110 The strategy proposed in this article aims to establish an OCPLS classification model and
111 characterize it by main qualitative and semi-quantitative performance parameters ($CC\alpha$, $CC\beta$,
112 and UR) estimated from PCC. Once characterized, in the prediction stage of new samples, the
113 main qualitative parameters give a global probability of wrong assignments while the semi-
114 quantitative parameters give similar information but at different concentration levels.

115 To minimize the probability of error (wrong assignment), in this work we propose defining two
116 class limits (lower and upper) as model distances, resulting in three distance regions: i) non-
117 adulterated region with distance values below the lower limit, ii) adulterated region with
118 distance values greater than the upper limit, and iii) uncertainty region with distance values
119 falling between both limits [40]. The assignment of unknown samples will be based on the region
120 where their calculated distance value falls. If a sample falls in the uncertainty region, it will be
121 considered inconclusive and should be submitted for further confirmatory analysis.

122 Numerous studies in food analysis have utilized multivariate qualitative methods to detect
123 potential adulteration. However, only a few of them have determined parameters related to the
124 limits of concentration detectable for the adulterant (Ruisánchez et al., 2022; Yang et al., 2022;
125 López et al., 2014; Fulgêncio et al., 2022; Pomerantsev et al., 2023). The strategy proposed in
126 the present work makes a noteworthy contribution in this sense, taking advantage of PCC to

127 determine the decision limit ($CC\alpha$) to evaluate the model before prediction. Setting two limits
128 eliminates the probability of error in the prediction of unknown samples but part of them may
129 be predicted as inconclusive.

130

131 **2. Experimental part**

132

133 *2.1. Samples*

134 Two batches of eucalyptus (EUC) and wild (W) types of honey were obtained directly from
135 traceable producers in Minas Gerais State, Brazil. Seven proportions of those batches were
136 prepared in five replicates to form 35 unadulterated samples for each type. Inverted sugar syrup
137 (ISS) was added in different quantities to the seven formulated batches of unadulterated
138 samples to obtain six levels of adulteration (1.0, 3.0, 9.0, 15.0, 21.0, and 27.0% w/w), resulting
139 in 42 adulterated samples.

140 All samples were weighed using a Shimadzu AUX 220 analytical balance with a calibrated scale
141 in 300 g flasks and in 15 mL or 50 mL Falcon tubes, followed by manual homogenization and
142 storage at room temperature (19°C to 25°C) until the moment of the analysis. In the case of a
143 honey sample crystallized, it was placed in a water bath at 39-40°C for 5 minutes.

144 *2.2. Instrumental measurements*

145 Time-domain NMR analysis was performed using the Minispec ND mq-20 from Bruker Biospin
146 GmbH (Rheinstetten, Germany), operating at a ^1H frequency of 20 MHz (0.47 T) with a 10 mm
147 probe head. The Transverse relaxation decays (T_2) were measured by a Carr-Purcell-Meiboom-
148 Gill (CPMG) pulse sequence with a 90° and 180° pulse length of 2.76 and 5.42, respectively, an
149 echo time of 500 μs , a recycle delay of 2 s, and 54 dB of gain. A homogenized sample portion
150 (approximately 5 mL) was placed on a crystal NMR tube and heated in a thermostatic bath at

151 40°C. The tube was introduced into the Minispec and analyzed with three readings of 16 scans
152 each, generating a relaxation curve of 500 data points for each sample.

153 2.3. Software

154 The data obtained were processed using MATLAB, version 8.0.0.783-R2012b (Natick, MA, USA),
155 and PLS Toolbox 7.0.2 (Eigenvector Research Inc., Wenatchee, WA, USA). The Matlab routine for
156 OCPLS was provided by Lu Xu (Xu et al., 2014). For the Inverse Laplace Transform the webapp
157 `rmn-ilt.streamlit.app` was used, developed by Tiago B. Moraes (Moraes et al.).

158

159 3. Theoretical background

160

161 3.1. One-class partial least squares classifier (OCPLS)

162 OCPLS classifier is based on partial least squares (PLS). For each sample, two statistical
163 parameters emerge once the number of latent variables (LV) is fixed. The first, Hotelling's T^2 ,
164 relies on the score distance (SD), quantifying the distance from an object to the center of the
165 class. The second is the absolute centered residual (ACR), which is a measure of dispersion or
166 residuals in the projection onto the vector of the OCPLS regression coefficients. A confidence
167 limit is established for both statistics, SD_{lim} and ACR_{lim} , that corresponds to the limit of the class
168 at a specified confidence level (typically 95%) (Xu et al., 2013; Fageerzada et al., 2020). A sample
169 must exhibit values for both statistical parameters below these limits to be classified as
170 belonging to a target class. Another criterion for sample assignment involves calculating the
171 reduced distance of a sample from its class ($d_{i,r}$) which is calculated according to the following
172 equation, Eq. (1). If that is the case, for sample prediction its distance value ($d_{i,r}$) should be
173 compared to a reduce distance class limit. Different distance values can be used as a limit to
174 assign a sample to a specific class, for instance, 1 (Miaw et al., 2018; Gondim et al., 2017), $\sqrt{2}$

175 (Bevilaqua et al., 2013; Durante et al., 2011), or an optimized distance obtained by applying ROC
176 curves (Ruisánchez et al., 2022). Usually, it is used the distance value established at 1, so $d_{i,r} \leq 1$.

177

$$178 \quad d_{i,r} = \sqrt{\left(\frac{SD_i}{SD_{lim}}\right)^2 + \left(\frac{ACR_i}{ACR_{lim}}\right)^2} \quad \text{Eq. (1)}$$

179

180 where SD_i and ACR_i are the statistical parameters of a sample “i” and the SD_{lim} and ACR_{lim} are the
181 corresponding statistical class limits at a determinate level of significance.

182 If the model is considered appropriate for the case under study, the next step is the prediction
183 of unknown samples. Instead of just using a distance class limit (usually $d_r \leq 1$), it is suggested to
184 define two distance class limits, the upper (d_{upper_lim}) and the lower (d_{lower_lim}) class limits. Recent
185 articles have also proposed the use of two decision limits (Rovira et al., 2023; Quintanilla-Casas
186 et al., 2020). In the sample prediction step, the sample is assigned according to $d_{i,r} \leq d_{lower_lim}$,
187 $d_{i,r} \geq d_{upper_lim}$, or in between both ($d_{lower_lim} \leq d_{i,r} \leq d_{upper_lim}$).

188

189 3.2. Semi-quantitative figures of merit

190 Sensitivity, specificity, and accuracy (López et al., 2015; Cuadros-Rodríguez et al., 2016; Ballabio
191 et al., 2018) are the main performance parameters for qualitative method validation. When
192 detecting food adulteration, these parameters can be related to the percentage of adulterant
193 present in the sample. This dependency can be observed by establishing PCC. PCCs have been
194 referred to by different names, such as the probability of detection (Wehling et al., 2011),
195 performance curves (Song et al., 2001), and probability of identification (LaBudde et al., 2012),
196 among others. PCC illustrates the probability (frequency) of obtaining a positive result (P(X))
197 against the concentration level of the analyte (Wehling et al., 2011; Macarthur et al., 2012). A

198 positive response will be obtained when the sample is predicted as adulterated or, in other
199 words, the sample does not belong to the model established for unadulterated/authentic
200 samples. So, when the adulterant is absent, the probability of obtaining a positive result should
201 be zero or close to zero ($P(X)=0\%$). But when the concentration of the adulterant increases, the
202 probability should increase until it reaches 100%. PCC is obtained according to Eq. 2, estimating
203 the parameters by minimizing the root mean square of the residuals (RMSE) (López et al., 2015).

204

$$205 \quad P(X) = \frac{a}{1 - e^{-(b+(c*x))}} + d \quad \text{Eq. (2)}$$

206 where $P(X)$ is the probability of a positive model outcome, x is the level/percentage of
207 adulterated present in the sample, and a , b , c and d are the parameters to be estimated.

208 From PCC, three additional semi-quantitative parameters can be obtained (López et al., 2015;
209 Gondim et al., 2017; Trullols et al., 2005):

- 210 - The decision limit ($CC\alpha$) is the minimum concentration of the analyte (adulterant) that
211 can be reliably detected or identified in a sample with a low statistical certainty, usually
212 5%. This value is obtained from the intersection of the PCC with a horizontal line at a
213 certain value of $P(X)$, usually 5%.
- 214 - The detection capability ($CC\beta$) is the concentration of the analyte (adulterant), from
215 which its presence can be reliably detected or identified in a sample with a statistical
216 certainty of 95%. The $CC\beta$ is obtained from the intersection of PCC with an upper
217 horizontal line at $P(X)= 95\%$.
- 218 - The unreliability region (UR) is the range between $CC\alpha$ and $CC\beta$, if a sample falls into UR,
219 it will be assigned as inconclusive and should undergo a confirmatory analysis.

220

221 *3.3. Receiver Operating Characteristic Curve*

222 The Receiver Operating Characteristic curve is a graph that illustrates the performance of a
223 classification model for a considered parameter, for example, the class distance limit. For each
224 value of the parameter studied, the curve plots sensitivity against 1-specificity (Ruisánchez et
225 al., 2021; Gondim et al., 2017).

226 The curve allows us to identify different points based on user interest. Normally it is used to
227 obtain the optimal value of the parameter under study which is considered as the point that
228 balances sensitivity and specificity (Ruisánchez et al., 2021). In this work, it is proposed to use
229 the ROC, to set the two class limits, which correspond to the class limit at which sensitivity
230 reaches 100%, regardless of specificity, and the class limit at which specificity reaches 100%,
231 regardless of sensitivity.

232 **4. Results and discussion**

233
234 Fig. 1 shows the relaxation curves of the non-adulterated EUC honey samples and adulterated
235 EUC honey with inverted sugar at different percentages (1-27%). The intensity of the signals of
236 mean curves of maximum normalized T^2 relaxation profiles obtained with the CPMG pulse
237 sequence increases with the level of adulteration (Fig. 1a). Similarly, Fig. 1b shows the inverse
238 Laplace transform (ILT) applied to these CPMG decays, where can be observed the distributed
239 T^2 relaxation time in the logarithmic scale for non-adulterated and adulterated samples. Non-
240 adulterated (dark green line) and adulterated samples at 1% (orange line) and 3% (blue line)
241 have relaxation water populations between 4.76 – 4.85 ms, from the start of their bands (Fig.
242 1b), as can be seen in Table 1. As the level of adulteration increases, these bands are shifted
243 reaching relaxation water populations between 6.85 – 7.37 ms (Table 1) for adulteration at 21%
244 (black line) and 27% (light green line). Therefore, the relaxation time depends on the percentage
245 of adulteration, indicating that water mobility was lower for non-adulterated than for
246 adulterated honey. Similar plots (Figure 1S and Table 1S) were obtained for W samples.

247 ----- Figure 1 -----

248 ----- Table 1 -----

249 The nature of these differences in Time-domain NMR (TD-NMR) signals allowed the problem to
250 be addressed through a semi-quantitative classification approach. Fig. 2 schematically shows a
251 flowchart of the steps involved in the development and validation of a multivariate classification
252 model. The first step was the selection of the proper classification method. As the aim was to
253 detect whether a sample is adulterated or not, a one-class classification model was chosen,
254 namely OCPLS [Xu, L. et al. 2014 and Moraes, T.B. et al., 2023].

255 ----- Figure 2 -----

256 Two independent one-class models were built for each of the two origins of honey samples, EUC
257 and W. Due to the reduced number of authentic samples, models were built with all 35 non-
258 adulterated samples. Four and three LVs were selected for EUC and W models respectively,
259 based on the minimum cross-validation classification error. Figs. 2S shows the OCPLS model
260 assignments plots of SD versus ACR obtained in the prediction for each type of honey,
261 considering the reduced SD and ACR values and class limit at $d_r \leq 1$.

262 Following the flowchart (Fig. 2), once the model is built, the next step is to validate it by
263 estimating its performance parameters, which gives a global performance of the model. In the
264 prediction of the non-adulterated EUC and W samples, considering the class limit at $d_r \leq 1$, the
265 sensitivity was 94.3%. This means that both models wrongly assign as adulterated 2 out of 35
266 samples when they are not. For the EUC and W models, 57.1% and 66.7% specificity were
267 achieved, meaning that 18 and 14 samples out of the 42 adulterated samples respectively, were
268 wrongly assigned as non-adulterated when indeed they were.

269 As stated, the specificity provides overall information about the prediction of the adulterated
270 samples. However, it does not provide information about the model behaviour for each level of

271 adulteration. When there are samples available at different adulteration levels, useful
272 information can be obtained related to the level of adulteration from PCC. Fig. 3 shows the PCC
273 estimated from OCPLS models for EUC (Fig. 3a) and W (Fig. 3b) honey samples. In addition to
274 authentic samples, 42 samples adulterated with inverted sugar syrup (7 samples for each level
275 of adulteration, between 1% and 27%) were used for testing each model. Table 2 presents the
276 equations and parameters related to fitted PCC.

277 ----- Figure 3 -----

278 ----- Table 2 -----

279 $CC\alpha$ values were obtained from the intersection of the PCC curve with the horizontal line fixed
280 at $P(X)$ equal to 5% and 10% for W and EUC honey PCC curves, respectively. From Fig. 3, $CC\alpha$ is
281 equal to 2% (w/w) for W honey and 5% (w/w) for EUC honey. Samples adulterated at
282 percentage levels below or close to $CC\alpha$ cannot be differentiated from the non-adulterated
283 samples. $CC\beta$ values were obtained from the intersection of the PCC curve with the horizontal
284 line fixed at a $P(X)=95\%$. $CC\beta$ was estimated at 12% (w/w) for both models. So, above 12% of
285 adulterant content, all samples will be assigned as adulterated with 95% certainty. In addition,
286 in both cases, above 15% of adulterant content, they can be detected with 100% of certainty.

287 It is important to emphasize that although α and β values are usually set at 0.05 (5%), the user
288 can define other values, which will change $CC\alpha$, $CC\beta$, and consequently the uncertainty interval.
289 In addition to $CC\alpha$, $CC\beta$, other probability values could be worth exploring. For instance, the
290 adulteration level at which there are 50% of properly obtaining a positive assignment (sample is
291 adulterated). From Fig. 3 (grey dotted lines), these values correspond to a percentage of
292 adulteration of 10% and 7% for EUC and W, respectively.

293 Once the PCCs have been evaluated, confirming that samples adulterated at 1% and 3% do not
294 differ from the non-adulterated samples, the sensitivity and specificity values have been

295 recalculated without them, also considering the class limit at $d_{i,r} \leq 1$. The sensitivity value remains
296 the same since the non-adulterated samples are consistent, resulting in a value of 94.3% for
297 both models. Specificity was 85.2% and 96.3% for the EUC and W models respectively.

298 Following the scheme in Fig. 2, if the model is considered appropriate for the case under study,
299 the next step is the prediction of unknown samples. Instead of just using the class limit, usually
300 established at $d_r \leq 1$, it is suggested to define two class limits, the upper limit (d_{upper_lim}) and the
301 lower limit (d_{lower_lim}). Both limits were established with 100% confidence obtained from the ROC
302 curves (Fig. 4). Fig. 4a shows the ROC curves from OCPLS models for EUC and Fig. 4b for W honey
303 samples. They have been built without considering adulterated samples with 1% and 3%
304 adulterant since the PCC curves showed that these cannot be differentiated from the non-
305 adulterated samples. The d_{upper_lim} was obtained from the maximum distance value at which
306 sensitivity is 100% (indicated on ROC plots with a red circle). Similarly, the d_{lower_lim} was obtained
307 from the minimum distance value at which the specificity is 100% (indicated on ROC plots with
308 a green circle). The d_{upper_lim} was set around 1.3 in both EUC model W model, while the d_{lower_lim}
309 was established at 0.68 and 0.61 for EUC and W models, respectively.

310 ----- Figure 4 -----

311 In Fig. 5, the two-class limits (d_{upper_lim} and d_{lower_lim}) for each type of honey are presented. It
312 should be noted that adulterated samples hardly appear in it because their distances are greater
313 than those in the graph (see Figures S2a and S2b). Table 3 presents the predictions of the
314 different types of samples depending on whether two limits or 1 limit (at a distance less than or
315 equal to 1) are considered.

316 ----- Figure 5 -----

317 ----- Table 3 -----

318 At the prediction stage, the percentage of non-adulterated samples that are correctly predicted
319 decreases considerably with the two-limit strategy, much more so in the case of W honey. In the
320 case of adulterated samples, the percentage of correctly predicted samples is similar in both
321 strategies for the two types of honey.

322 However, as can be seen in the table, in the case of using a single limit, the prediction of the
323 samples carries an associated probability of error, for example, there is a 5.7% probability of
324 wrongly predicting a non-adulterated sample as adulterated. In the case of following the two-
325 limit strategy if a sample is predicted as non-adulterated, there is a 100% probability that it is.
326 The counterpart of not having an error in the assignment is that 17.1% of the non-adulterated
327 samples would require considering a confirmatory analysis

328 The main advantage of establishing two class limits is eliminating the probability of error
329 because the two limits permit no assignment ambiguity. Below the lower threshold are only
330 samples assigned as non-adulterated and above the upper threshold there are only ones
331 assigned as adulterated.

332 From the practical point of view, end-users have the flexibility to define the UR limits based on
333 the specific nature of the problem. For instance, in the current study, if the end-user is aware
334 that honey samples are not expected to be adulterated below 10%, adjusting the d_{lower_lim} to
335 slightly higher values can effectively reduce the UR. Defining the two limits involves striking a
336 balance between the percentage of samples requiring confirmatory analysis and the acceptable
337 error percentage for the end user.

338

339 **5. Conclusions**

340 Two one-class models using OCPLS were developed to detect eucalyptus and wild honey
341 adulteration with inverted sugar syrup. These models were validated by establishing main

342 performance and semi-quantitative parameters ($CC\alpha$, $CC\beta$, and UR) obtained from performance
343 characteristic curves (PCC). The developed strategy allowed the examination of the model's
344 behaviour across different adulteration levels before the prediction step. The estimate of $CC\alpha$
345 provided insights into which/how many samples at the lowest levels of adulteration cannot be
346 discriminated from non-adulterated/authentic samples. The estimate of $CC\beta$ indicated the
347 adulteration level from which samples will be identified as adulterated with certainty.

348 The proposed strategy of defining two class limits allowed the establishment of an uncertainty
349 region, which implies determining with 100% certainty whether a sample is compliant or non-
350 compliant. Consequently, the probability of error in the prediction of unknown samples is
351 eliminated when compared to using only a single. For both types of honey, all samples
352 adulterated at the highest levels (15%, 21%, and 27% w/w) were correctly assigned as
353 adulterated, while all samples at 9% w/w were classified in the uncertainty region, thus
354 demanding to be subjected to confirmatory analysis. The approach developed in this article can
355 be easily adapted to the detection of other types of adulterants in honey.

356

357 **Acknowledgments**

358 The authors would like to thank the Brazilian government agencies CNPq, FAPEMIG, and CAPES
359 for financial support. FAPEMIG (Fundação de Amparo à Pesquisa do Estado de Minas Gerais) is
360 particularly acknowledged for the financial support of the project "Rede Mineira de Ciências
361 Forenses" (RED00042-16) and the project "Programa de Cooperação Acadêmica em Segurança
362 Pública e Ciências Forenses" (Edital CAPES 16/2020). The authors also would like to thank
363 Professor Ricardo Geraldo de Sousa, coordinator of the Polymer Science and Technology
364 Laboratory (LCTP) at UFMG, for the lend of low-field nuclear magnetic resonance equipment.

365 This research study is part of a project supported through joint funding under the collaborative
366 agreement between the Diputació de Tarragona and the Universitat Rovira I Virgili, covering the

367 period from 2020 to 2023. This project was carried out in the year 2023, with the reference
368 2023PIN-DIPTA-URV01 and focuses on the “Training of pre-doctoral research staff”.

369

370 **References**

371 Antônio, D. C., de Assis, D. C. S., Botelho, B. G., Sena, M. M. (2022). Detection of adulterations
372 in a valuable Brazilian honey by using spectrofluorimetry and multiway classification. *Food*
373 *Chemistry*, 370, 131064. <https://doi.org/10.1016/j.foodchem.2021.131064>.

374 AOAC, C-4 plant sugars in honey, Internal Standard Stable Carbon Isotope Ratio Method.
375 Association of Official Analytical Chemists 1998 p.4.

376 Ballabio, D., Grisoni, F., Todeschini, R. (2018). Multivariate comparison of classification
377 performance measures. *Chemometrics and Intelligent Laboratory Systems*, 174, 33–44. <https://doi.org/10.1016/j.chemolab.2017.12.004>.

379 Başar, B., Özdemir, D. (2018). Determination of honey adulteration with beet sugar and corn
380 syrup using infrared spectroscopy and genetic-algorithm-based multivariate calibration. *Journal*
381 *of Science of Food and Agriculture*, 98, 5616-5624. <https://doi.org/10.1002/jsfa.9105>.

382 Belton, P. S. (1990). Can nuclear magnetic resonance give useful information about the state of
383 water foodstuffs? *Agricultural Food Chemistry*, 2, 179-209.

384 Bertram, H. C., Karlsson, A. H., Rasmussen, M., Pedersen, O. D., Døntup, S., Andersen, H. J.
385 (2001). Origin of multiexponential T₂ relaxation in muscle myowater. *Journal of Agricultural and*
386 *Food Chemistry*, 49, 3092-3100. <https://doi.org/10.1021/jf001402t>.

387 Bevilaqua, M., Bucci, R., Magrì, A.D., Nescatelli, R., Marini, F. (2013). Classification and class-
388 modeling in: F. Marini (Editor). Data handling in science and Technology, *Elsevier*, 28, 171-233.
389 <https://doi.org/10.1016/B978-0-444-59528-7.00005-3>.

390 Biaswas, A., Chaudhari, S. R. (2024). Exploring the role of NIR spectroscopy in quantifying and
391 verifying honey authenticity: A review. *Food Chemistry*, 445, 138712.
392 <https://doi.org/10.1016/j.foodchem.2024.138712>.

393 Biaswas, A., KS, N., Jaygadkar, S. S., Chaudhari, S. R. (2023). Enabling honey quality and
394 authenticity with NMR and LC-IRMS based platform. *Food Chemistry*, 416, 135825.
395 <https://doi.org/10.1016/j.foodchem.2023.135825>.

396 Brar, D.S., Pant, K., Krishnan, R., Kaur, S., Rasane, P., Nanda, V., Saxena, S., Gautam, S. (2023). A
397 comprehensive review on unethical honey: Validation by emerging techniques. *Food Control*,
398 145, 109482. <https://doi.org/10.1016/j.foodcont.2022.109482>.

399 Brereton, R. G., Lloyd, G. R. (2014). Partial least squares discriminant analysis: taking the magic
400 away. *Journal of Chemometrics*, 28, 213-225. <https://doi.org/10.1002/cem.2609>.

401 Cárdenas-Escudero, J., Galán-Madruga, D., Cáceres, J. O. (2023). Rapid, reliable and easy-to-
402 perform *chemometric-less* method for rice syrup adulterated honey detection using FITR-ATR.
403 *Talanta*, 253, 123961. <https://doi.org/10.1016/j.talanta.2022.123961>.

404 Cengiz, O. (2023). Classification of Edible oils by using Time Domain NMR (TD-NMR) technique
405 and Microwave (MW) dielectric spectroscopy. *Food Analytical Methods*, 16, 1529-1536.
406 <https://doi.org/10.1007/s12161-023-02520-6>.

407 Ciursă, P., Pauliuc, D., Dranca, F., Ropciuc, S., Oroian, M. (2021). Detection of honey adulterated
408 with agave, corn, inverted sugar maple and rice syrups using FTIR analysis. *Food Control*, 130,
409 108266. <https://doi.org/10.1016/j.foodcont.2021.108266>.

410 Codex Alimentarius Commission (2001). Revised codex standard for honey. Alinorm, 19–26, Food
411 and Agriculture Organization of the United Nations, Rome.

412 Coimbra, P. T., Bathzar, C. F., Guimarães, J. T., Coutinho, N. M., Pimentel, T. C., Neto, R. P. C.,
413 Esmerino, E. A., Freitas, M. Q., Silva, M. C., Tavares, M. I. B., Cruz, A. G. (2020). Detection of

414 formaldehyde in raw milk by time domain nuclear magnetic resonance and chemometrics. *Food*
415 *Control*, 110, 107006. <https://doi.org/10.1016/j.food.cont.2019.107006>.

416 Colnago, L. A., Wiesman, Z., Pages, G., Musse, M., Monaretto, T., Windt, C. W., Rondeau-Moro,
417 C. (2021). Low field, time domain NMR in the agriculture and agrifood sectors: An overview of
418 applications in plants, food and biofuels. *Journal of Magnetic Resonance*, 323, 106899.
419 <https://doi.org/10.1016/j.jmr.2020.106899>.

420 Corps, A.I., Rodriguez, N., Guzman, F.J., Rodriguez, R.C., Rios, A. (2021). Screening-confirmation
421 strategy for nanomaterials involving spectroscopic analytical techniques and its application to
422 the control of silver nanoparticles in pastry samples. *Spectrochimica Acta A*, 246, 119015.
423 <https://doi.org/10.1016/j.saa.2020.119015>.

424 Cuadros-Rodríguez, L., Pérez-Castaño, E., Ruiz-Samblá, C. (2016). Quality performance metrics
425 in multivariate classification methods for qualitative analysis. *TrAC Trends in Analytical*
426 *Chemistry*, 80, 612-624. <https://doi.org/10.1016/j.trac.2016.04.021>.

427 Curti, E., Anedda, R. (2023). TD-NMR as a quality control tool for dairy products: a study on Fiore
428 Sardo PDO cheese. *Food and Bioprocess Technology*, 16, 459-465.
429 <https://doi.org/10.1007/s11947-022-02947-5>.

430 De Souza, R. R., Fernandes, D. D. d.S., Diniz, P. H. G. D. (2021). Honey authentication in terms of
431 its adulteration with sugar syrups using UV-Vis spectroscopy and one-class classifiers. *Food*
432 *Chemistry*, 365, 130467. <https://doi.org/10.1016/j.foodchem.2021.130467>.

433 Dias, L. G., Bruni, A. R. S., Anizelli, C. P., Zangirolami, M., Lima, P. C., Estevinho, L. M., Bona, E.
434 (2022). Semi-quantitative discrimination of honey adulterated with cane sugar solution by an
435 ETongue. *Chemistry & Biodiversity*, 19, e2022006. <https://doi.org/10.1002/cbdv.202200698>.

436 Durante, C., Bro, R., Cocchi, M. (2011). A classification tool for N-way array based on SIMCA
437 methodology. *Chemometrics and Intelligent Laboratory Systems*, 106, 73–85. [https://doi.org/](https://doi.org/10.1016/j.chemolab.2010.09.004)
438 10.1016/j.chemolab.2010.09.004.

439 Egido, C., Saurina, J., Sentellas, S., Núñez, O. (2024). Honey fraud detection based on sugar syrup
440 adulterations by HPLC-UV fingerprinting and chemometrics. *Food Chemistry*, 436, 137758.
441 <https://doi.org/10.1016/j.foodchem.2023.137758>.

442 European Commission. (2018). Meeting Report of Technical Round Table on Honey
443 Authentication. Retrieved from [https://ec.europa.eu/jrc/sites/jrcsh/files/](https://ec.europa.eu/jrc/sites/jrcsh/files/ares1815690741_technical_round_table_on_honey_adulteration_report.pdf)
444 ares1815690741_technical_round_table_on_honey_adulteration_report.pdf.

445 Fageerzada, M. A., Lohumi, S., Joshi, R., Kim, M. S., Baek, I., Cho, B-K. (2020). Non-targeted
446 detection of adulterants in almond powder using spectroscopic techniques combined with
447 chemometrics. *Foods*, 7, 876. <https://doi.org/10.3390/foods9070876>.

448 Fakhlaei, R., Selamat, J., Khatib, A., Razis, A. F. A., Sukor, R., Ahmad, S., Babadi, A. A. (2020). The
449 toxic impact of honey adulteration: A review. *Foods*, 9, 1538.
450 <https://doi.org/10.3390/foods9111538>.

451 Finch, E. D., Harmon, J. D., Muller, B. H. (1971). Pulsed NMR measurements of the diffusion
452 constant of water in muscle. *Archives of Biochemistry and Biophysics*, 147, 299-310.
453 [https://doi.org/10.1016/0003-9861\(71\)90337-7](https://doi.org/10.1016/0003-9861(71)90337-7).

454 Fulgêncio, A. C. C., Resende, G. A. P., Teixeira, M. C. F., Botelho, B. G., Sena, M. M. (2022).
455 Screening method for the rapid detection of diethylene glycol in beer based on chemometrics
456 and portable near-infrared spectroscopy. *Food Chemistry*, 391, 133258.
457 <https://doi.org/10.1016/j.foodchem.2022.133258>.

458 Gondim, C. de S., Junqueira, R. G., de Souza, S. V. C., Callao, M. P., Ruisánchez, I. (2017).
459 Determining performance parameters in qualitative multivariate methods using probability of

460 detection (POD) curves. Case study: two common milk adulterants. *Talanta*, 168, 23-30.
461 <https://doi.org/10.1016/j.talanta.2016.12.065>.

462 Gondim, C. S., Junqueira, R. G., de Souza, S. V. C., Ruisánchez, I., Callao, M. P. (2017). Detection
463 of several common adulterants in raw milk by MID-infrared spectroscopy and one-class and
464 multi-class multivariate strategies. *Food Chemistry*, 230, 68–75.
465 <https://doi.org/10.1016/j.foodchem.2017.03.022>.

466 Guelpa, A., Marini, F., Plessis, A., Slabbert, R., Manley, M. (2017). Verification of authenticity and
467 fraud detection in South African honey using NIR spectroscopy. *Food Control*, 73, 1388-1396.
468 <https://doi.org/10.1016/j.foodcont.2016.11.002>.

469 Huang, F., Song, H., Guo, L., Guang, P., Yang, X., Li, L., Zhao, H., Yang, M. (2020). Detection of
470 adulteration in Chinese honey using NIR and ATR-FTIR spectral data fusion. *Spectrochimica Acta*
471 *A*, 235, 118297. <https://doi.org/10.1016/j.saa.2020.118297>.

472 Instrução Normativa Nº 11. (2000). Regulamento técnico de identidade e qualidade do mel.
473 *Ministério da Agricultura, Pecuária e Abastecimento*. Available at:
474 [https://www.gov.br/agricultura/pt-br/assuntos/defesa-agropecuaria/suasa/regulamentos-](https://www.gov.br/agricultura/pt-br/assuntos/defesa-agropecuaria/suasa/regulamentos-tecnicos-de-identidade-e-qualidade-de-produtos-de-origem-animal-1/IN11de2000.pdf)
475 [tecnicos-de-identidade-e-qualidade-de-produtos-de-origem-animal-1/IN11de2000.pdf](https://www.gov.br/agricultura/pt-br/assuntos/defesa-agropecuaria/suasa/regulamentos-tecnicos-de-identidade-e-qualidade-de-produtos-de-origem-animal-1/IN11de2000.pdf).

476 Jiménez-Carvelo, A. M., Cuadros-Rodríguez, L. (2020). The occurrence: a meaningful parameter
477 to be considered in the validation of multivariate classification-based screening methods-
478 application for authenticating virgin olive oil. *Talanta*, 208, 120467.
479 <https://doi.org/10.1016/j.talanta.2019.120467>.

480 LaBudde, R. A., Harnly, J. M. (2012). Probability of identification: a statistical model for the
481 validation of qualitative botanical identification methods, *Journal of AOAC International*, 95,
482 273-285. <https://doi.org/10.5740/jaoacint.11-266>.

483 Li, Q., Zeng, J., Lin, L., Zhang, J., Zhu, J., Yao, L., Wu, Z. (2020). Low risk of category misdiagnosis
484 of rice syrup adulteration in three botanical origin honey by ATR-FTIR and general model. *Food*
485 *Chemistry*, 332, 127356. <https://doi.org/10.1016/j.foodchem.2020.127356>.

486 Li, S., Zhang, X., Shan, Y., Su, D., Ma, Q., Wen, R., Li, J. (2017). Qualitative and quantitative
487 detection of honey adulterated with high-fructose corn syrup and maltose syrup by using near-
488 infrared spectroscopy. *Food Chemistry*, 2018, 231-236.
489 <https://doi.org/10.1016/j.foodchem.2016.08.105>.

490 Limm, W., Karunathilaka, S. R., Mossoba, M. M. (2023). Fourier transform infrared spectroscopy
491 and chemometrics for the rapid screening of economically motivated adulteration of honey
492 spiked with corn or rice syrup. *Journal of Food Protection*, 86, 100054.
493 <https://doi.org/10.1016/j.jfp.2023.100054>.

494 López, M. I., Callao, M. P., Ruisánchez, I. (2015). A tutorial on the validation of qualitative
495 methods: from the univariate to the multivariate approach. *Analytica Chimica Acta*, 891, 62-72.
496 <https://doi.org/10.1016/j.aca.2015.06.032>.

497 López, M. I., Colomer, N., Ruisánchez, I., Callao, M. P. (2014). Validation of multivariate screening
498 methodology. Case study: Detection of food fraud. *Analytica Chimica Acta*, 827, 28-33.
499 <https://doi.org/10.1016/j.aca.2014.04.019>.

500 Macarthur, R., von Holst, C. (2012). A protocol for the validation of qualitative methods of
501 detection. *Analytical Methods*, 4, 2744-2754. <https://doi.org/10.1039/C2AY05719K>.

502 Machado, G. O., Teixeira, G. G., Garcia, R. H. S., Moraes, T. B., Bona, E., Santos, P. M., Colagno,
503 L. A. (2022). Non-invasive method to predict the composition of requeijão cremoso directly in
504 commercial packages using time domain NMR relaxometry and chemometrics. *Molecules*, 27,
505 4434. <https://doi.org/10.3390/molecules27144434>.

506 Miaw, C. S. M., Sena, M. M., de Souza, S. V. C., Callao, M. P., Ruisánchez, I. (2018). Detection of
507 adulterants in grape nectars by attenuated total reflectance Fourier-transform mid-infrared
508 spectroscopy and multivariate classification. *Food Chemistry*, 266, 254–261.
509 <https://doi.org/10.1016/j.foodchem.2018.06.006>.

510 Moraes, T. B., Mazzero, L. P., Mendes, W. S. Inverse Laplace Transform WebApp. Available at:
511 <https://rmn-ilt.streamlit.app/> (visited last time: 16 nov. 23).

512 Orioian, M., Ropciuc, S., Paduret, S. (2018). Honey adulteration detection using Raman
513 spectroscopy. *Food Analytical Methods*, 11, 959-968. [https://doi.org/10.1007/s12161-017-](https://doi.org/10.1007/s12161-017-1072-2)
514 1072-2.

515 Pomerantsev, A. L., Vtyurina, D. N., Rodionova, O. Y. (2023). Limit of detection in qualitative
516 analysis: Classification Analytical Signal approach. *Microchemical Journal*, 195, 109490.
517 <https://doi.org/10.1016/j.microc.2023.109490>.

518 Quintanilla-Casas, B., Bustamante, J., Guardiola, F., García-González, D. L., Barbieri, S., Bendini,
519 A., Toschi, T.G., Vichi, S., Tres, A. (2020). Virgin olive oil volatile fingerprint and chemometrics:
520 towards an instrumental screening tool to grade the sensory quality. *LWT Food Science and*
521 *Technology*, 121, 108936. <https://doi.org/10.1016/j.lwt.2019.108936>.

522 Rachineni, K., Kakita, V. M. R., Awasthi, N. P., Shirke, V. S., Hosur, R. V., Shukla, S. C. (2022).
523 Identifying type of sugar adulterants in honey: Combined application of NMR spectroscopy and
524 supervised machine learning classification. *Current Research in Food Science*, 5, 272-277.
525 <https://doi.org/10.1016/j.crfs.2022.01.008>.

526 Ribeiro, R. O. R., Mársico, E. T., Carneiro, C. S., Monteiro, M. L. G., Júnior, C. C., Jesus, E. F. O.
527 (2014). Detection of honey adulteration of high fructose corn syrup by Low Field Nuclear
528 Magnetic Resonance (^1H NMR). *Journal of Food Engineering*, 135, 39-43.
529 <https://doi.org/10.1016/j.jfoodeng.2014.03.009>.

530 Ribeiro, U., Queiroz, L., Marassi, A., Carvalho, A., Barros, G., Consalter, D., Bezerra, J., Santos, A.,
531 Colnago, L.A., Machado, M. (2021). Development of a TD-NMR method to monitor Brazil nuts
532 oil content: a green and low-cost based approach. *Journal of Brazilian Chemistry Society*, 32,
533 1405-1412. <https://doi.org/10.21577/0103-5053.20210039>.

534 Ricardo, A. I. C., García, S. A., Bernardo, F. J. G., Ríos, A., Martín-Doimeadios, R. C. R. (2021).
535 Rapid assessment of silver nanoparticle migration from food containers into food simulants
536 using a qualitative method. *Food Chemistry*, 361, 130091.
537 <https://doi.org/10.1016/j.foodchem.2021.130091>.

538 Rovira, G., Miaw, C. S. W., Martins, M. L. C., Sena, M. M., de Souza, S. V. C., Callao, M. P.,
539 Ruisánchez, I. (2023). One-class model with two decision thresholds for the rapid detection of
540 cashew nuts adulteration by other nuts. *Talanta*, 253, 123916.
541 <https://doi.org/10.1016/j.talanta.2022.123916>.

542 Ruisánchez, I., Jiménez-Carvelo, A. M., Callao, M. P. (2021). ROC curves for the optimization of
543 one-class model parameters. A case study: authenticating extra virgin olive oil from a Catalan
544 protected designation of origin. *Talanta*, 222, 121564.
545 <https://doi.org/10.1016/j.talanta.2020.121564>.

546 Ruisánchez, I., Rovira, G., Callao, M. P. (2022). Multivariate qualitative methodology for semi-
547 quantitative information. A case study: adulteration of olive oil with sunflower oil. *Analytica*
548 *Chimica Acta*, 1206, 339785. <https://doi.org/10.1016/j.aca.2022.339785>.

549 Sahland, M., Karwita, S., Gozan, M., Hermansyah, H., Yohda, M., Yoo, Y. J., Pratami, D. K. (2019).
550 Identification and classification of hone's authenticity by attenuated total reflectance Fourier-
551 transform infrared spectroscopy and chemometric method. *Veterinary World*, 12, 1304-1310.
552 <https://doi.org/10.14202/vetworld.2019.1304-1310>.

553 Santos, P. M., Pereira-Filho, E. R., Colagno, L. A. (2016). Detection and quantification of milk
554 adulteration using time domain nuclear magnetic resonance (TD-NMR). *Microchemical Journal*,
555 124, 15-19. <https://doi.org/10.1016/j.microc.2015.07.013>.

556 Se, K. W., Wahab, R. A., Yaacob, S. N. S., Ghoshal, S.K. (2019). Detection techniques for
557 adulterants in honey: Challenges and recent trends. *Journal of Food Composition and Analysis*,
558 80, 16-32. <https://doi.org/10.1016/j.jfca.2019.04.001>.

559 Song, R., Schlecht, P. C., Ashley, K. (2001). Field screening test methods: performance criteria
560 and performance characteristics. *Journal of Hazardous Materials*, 83, 29-39.
561 [https://doi.org/10.1016/S0304-3894\(00\)00325-3](https://doi.org/10.1016/S0304-3894(00)00325-3).

562 Trullols, E., Ruisánchez, I., Rius, F.X., Huguet, J. (2005). Validation of qualitative methods of
563 analysis that use control samples. *TrAC Trends in Analytical Chemistry*, 24, 516-524.
564 <https://doi.org/10.1016/j.trac.2005.04.001>.

565 Wehling, P., LaBudde, R. A., Brunelle, S. L., Nelson, M. T. (2011). Probability of detection (POD)
566 as a statistical model for the validation of qualitative methods. *Journal of AOAC International*,
567 94, 335-347. <https://doi.org/10.1093/jaoac/94.1.335>.

568 Xu, L., Goodarzi, M., Shi, W., Cai, C.B., Jiang, J. H. (2014). A MATLAB toolbox for class modeling
569 using one-class partial least squares (OCPLS) classifiers. *Chemometrics and Intelligent Laboratory*
570 *Systems*, 139, 58-63. <https://doi.org/10.1016/j.chemolab.2014.09.005>.

571 Xu, L., Yan, S.M., Cai, C.B., Yu, X.P. (2013). One-class partial least squares (OCPLS) classifier.
572 *Chemometrics and Intelligent Laboratory Systems*, 126, 1-5.
573 <https://www.doi.org/10.1016/j.chemolab.2013.04.008>.

574 Yang, Q., Lin, H., Ma, J., Chen, N., Zhao, C., Guo, D., Niu, B., Zhao, Z., Deng, X., Chen, Q. (2022).
575 An improved POD model for fast semi-quantitative analysis of carbendazim in fruit by surface

576 enhanced raman spectroscopy. *Molecules*, 27, 4230.

577 <https://doi.org/10.3390/molecules27134230>.

578

Preprint not peer reviewed

Figure Captions

Fig. 1. (a) Maximum normalized T^2 mean relaxation curves obtained with CPMG pulse sequence for unadulterated/authentic and adulterated (1-27% w/w) eucalyptus honey samples. **(b)** T^2 relaxation spectra of the respective honey samples obtained via ILT. Color codes: dark green for non-adulterated honey, orange for samples adulterated at 1%, blue for adulterated at 3%, red for adulterated at 9%, purple for adulterated at 15%, dark blue for adulterated at 21% and light green for adulterated at 27%.

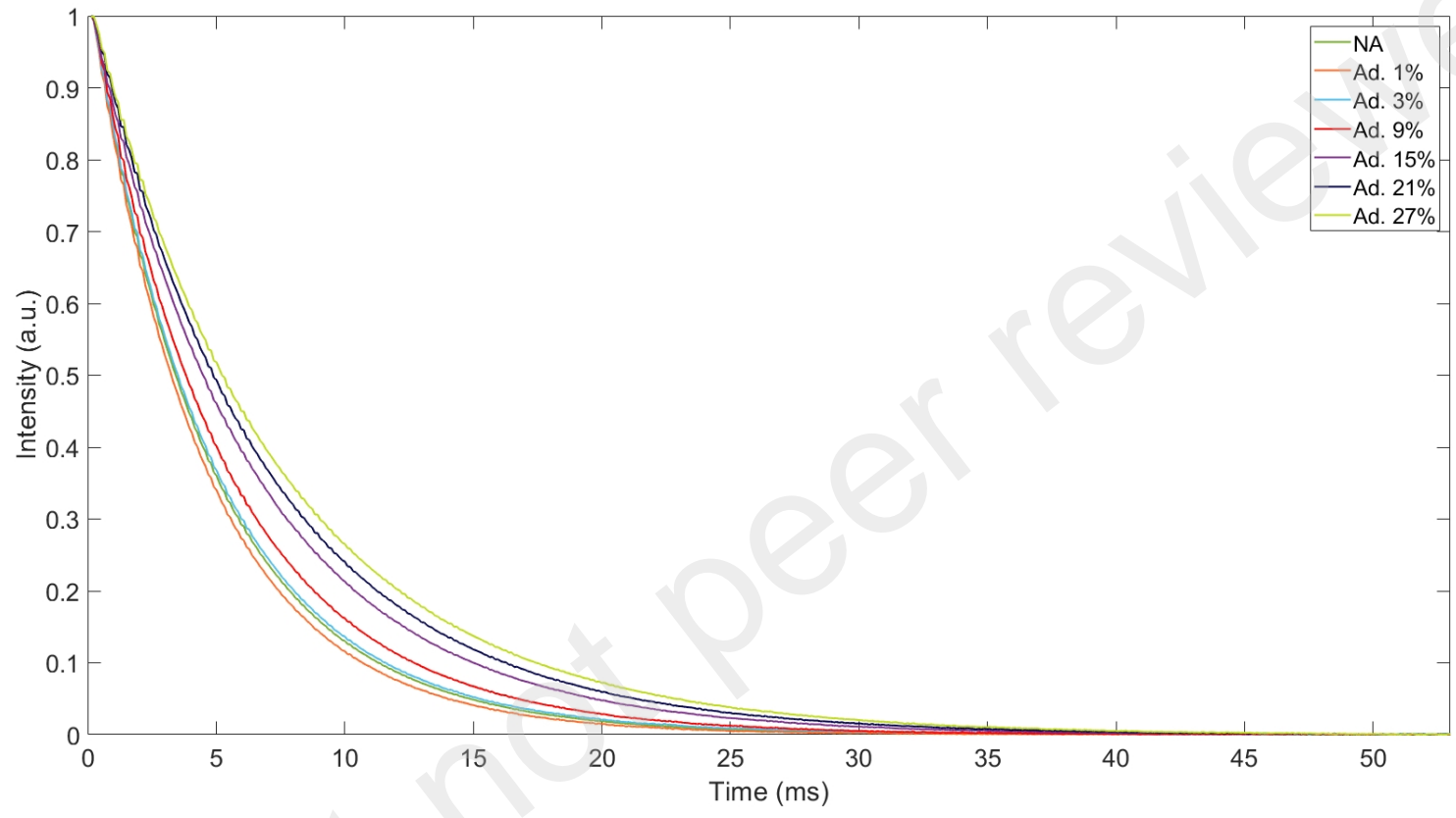
Fig. 2. Schematic flowchart describing the semi-quantitative chemometric approach developed in this work.

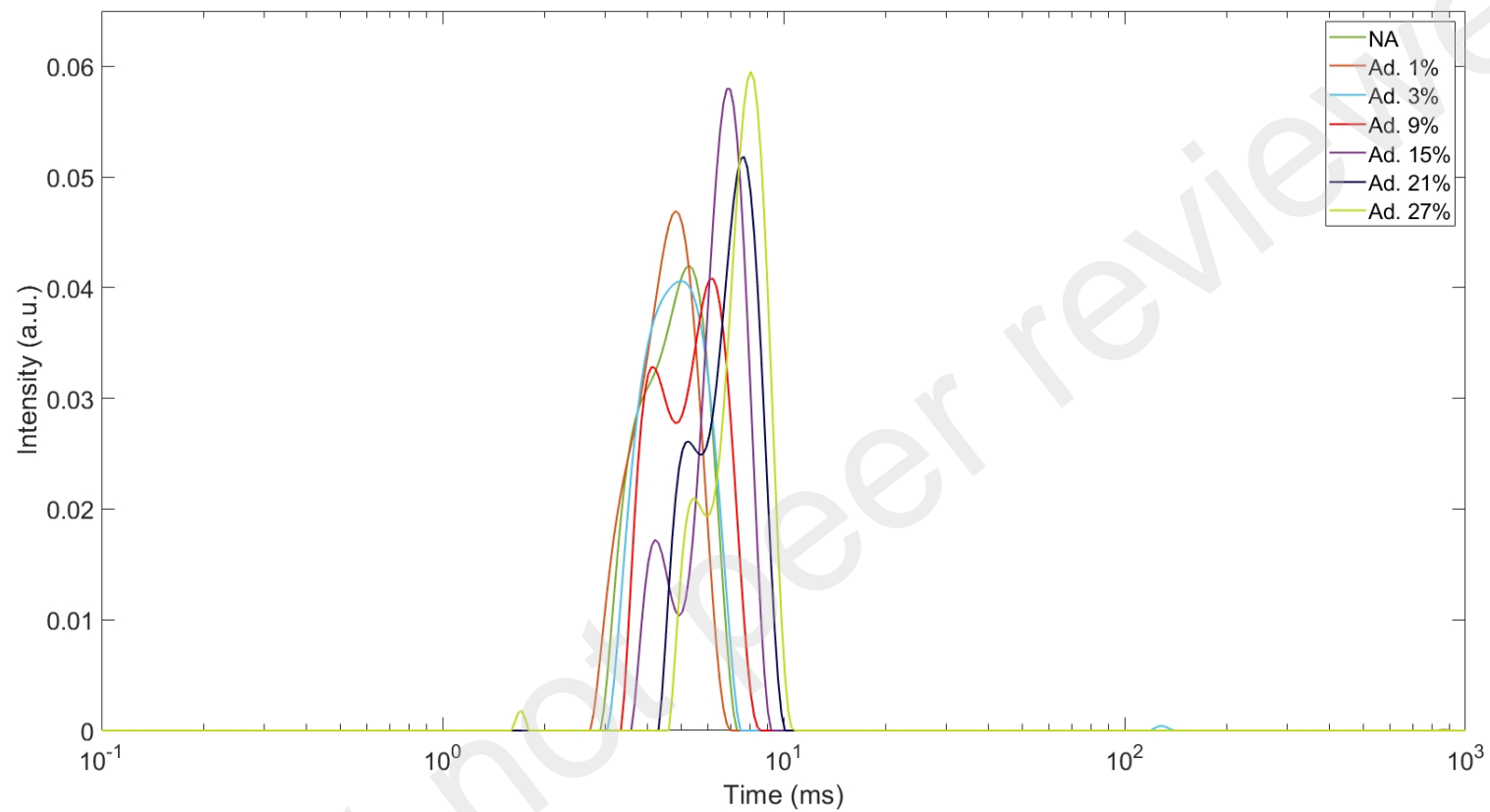
Fig. 3. Prediction results obtained by OCPLS models for detecting adulteration with inverted sugar syrup in **(a)** eucalyptus and **(b)** wild honey. Dashed lines indicate significance levels of 0.05 for both score distances (SD) and centered model residuals (ACR). Boxes inside the plots represent a zoom view of the acceptance region (SD and ACR values below the limits). Color and symbol codes: Full green circles for non-adulterated honey, empty orange triangles for samples adulterated at 1%, empty light blue triangles for adulterated at 3%, empty red triangles for adulterated at 9%, empty purple triangles for adulterated at 15%, empty dark blue triangles for adulterated at 21% and empty light green triangles for adulterated at 27%.

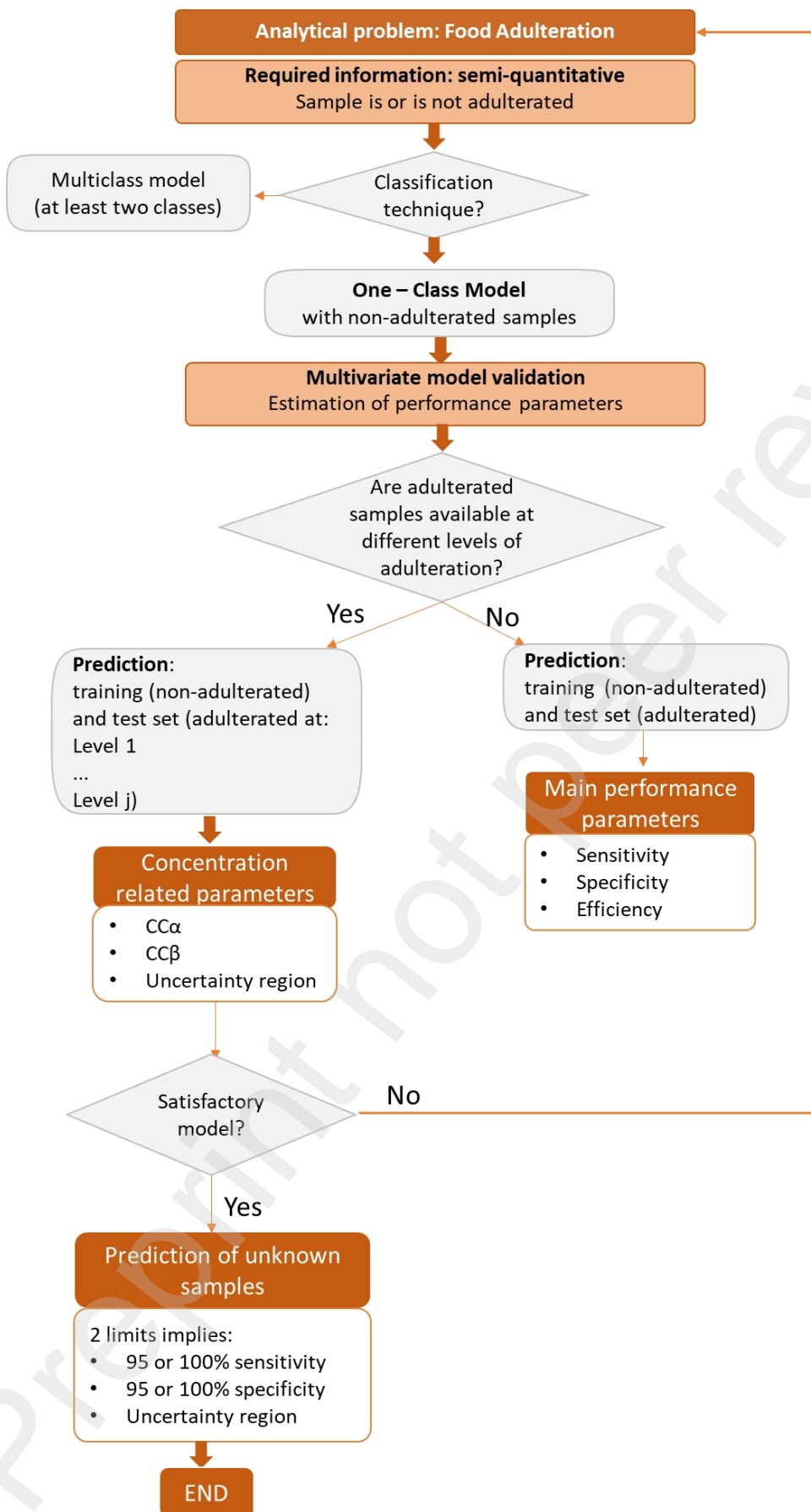
Fig. 4. Performance characteristic curves (PCC) constructed for estimating $CC\alpha$ and $CC\beta$ for OCPLS models built with adulterated **(a)** eucalyptus and **(b)** wild honey samples.

Fig. 5. Distances of all the analyzed samples to the OCPLS model. Vertical red and green solid lines represent the lower and upper limits estimated for the uncertainty region, respectively.

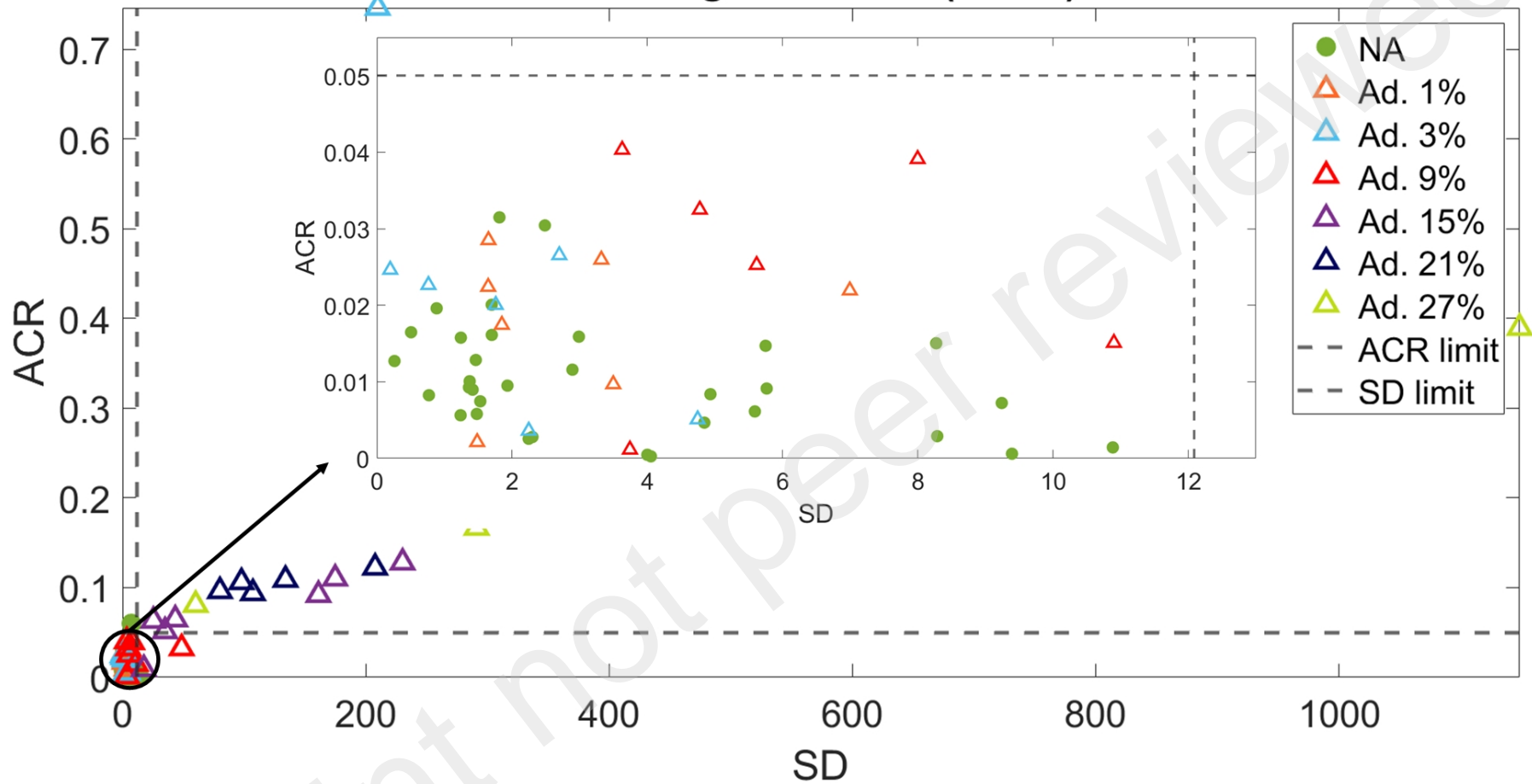
Color and symbol codes: Full green circles for non-adulterated honey, empty orange triangles for samples adulterated at 1%, empty light blue triangles for adulterated at 3%, empty red triangles for adulterated at 9%, empty purple triangles for adulterated at 15%, empty dark blue triangles for adulterated at 21% and empty light green triangles for adulterated at 27%.



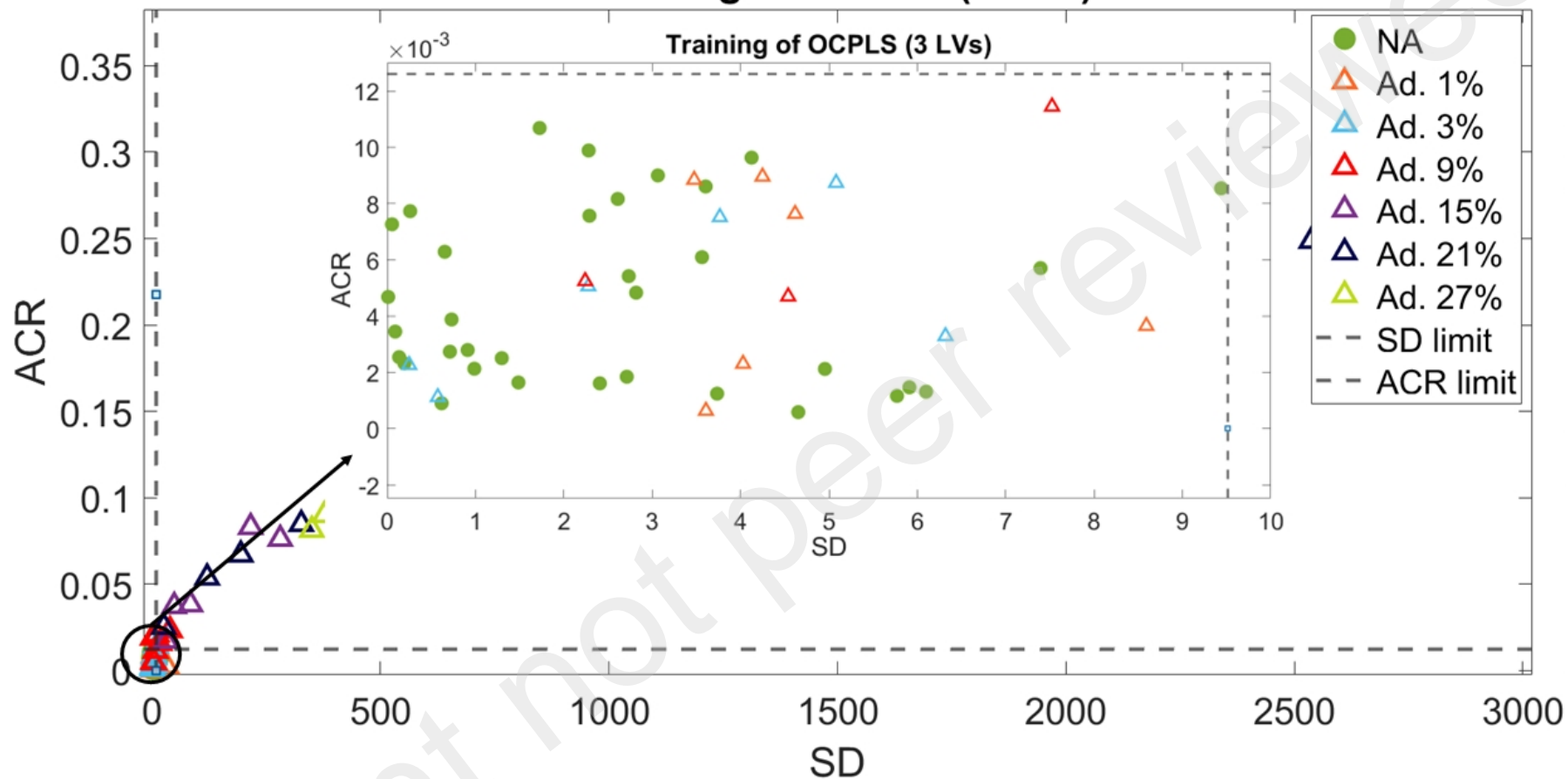


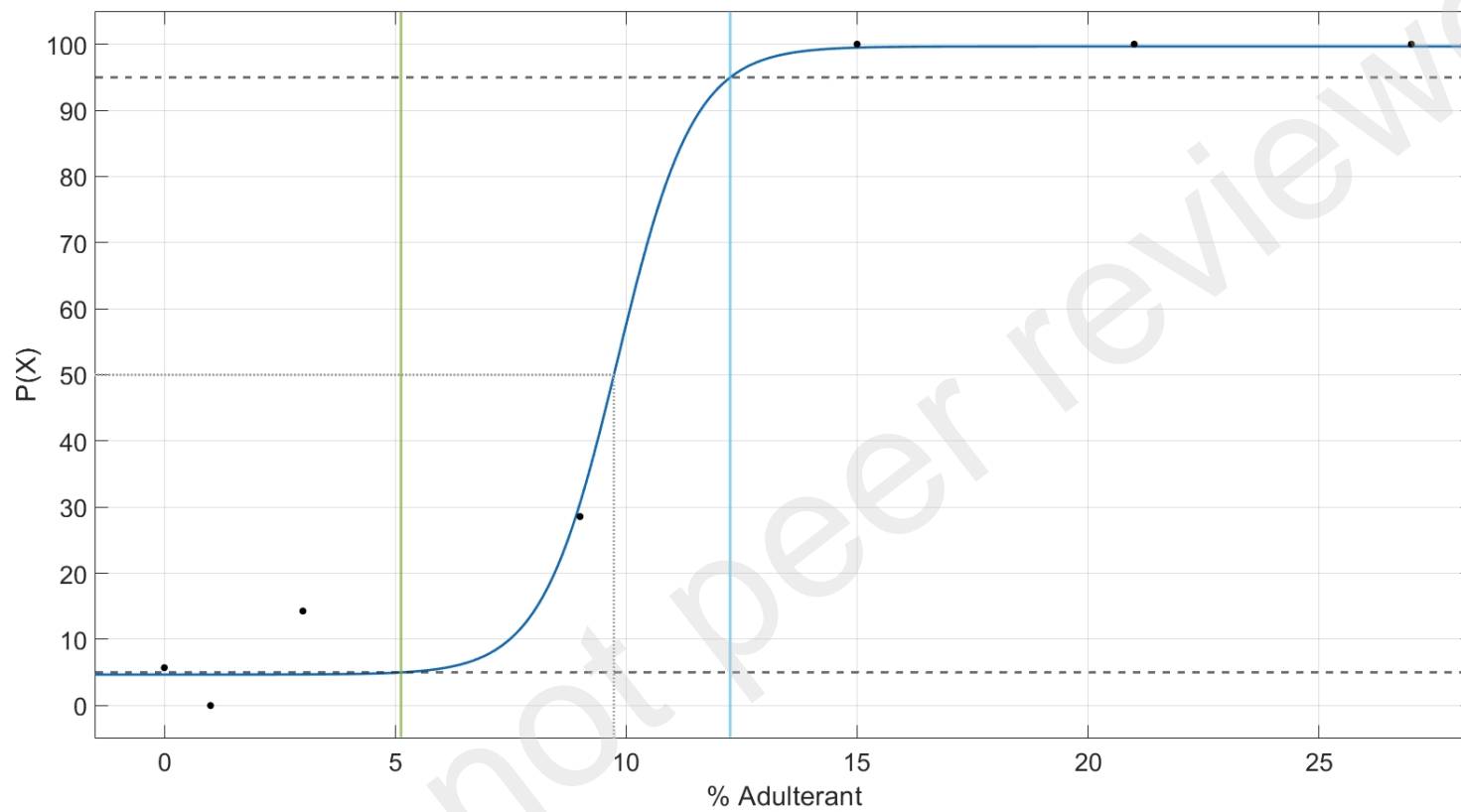


Training of OCPLS (4 LVs)



Training of OCPLS (3 LVs)





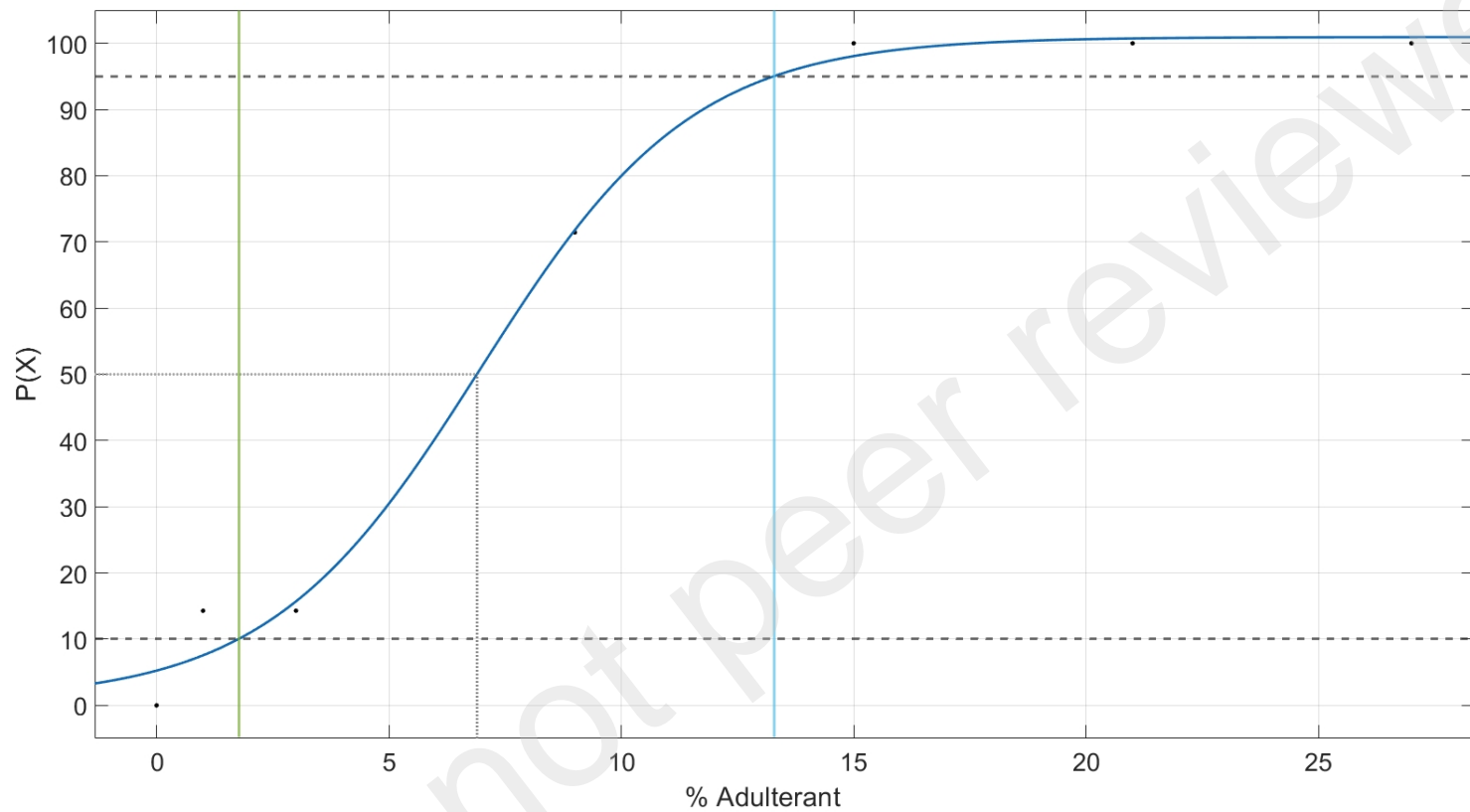


Table 1. T² parameters obtained in EUC honey according to level of adulteration with ISS

| Adulteration level | T² (ms) |
|---------------------------|---------------------------|
| 0% | 4.76 |
| 1% | 4.50 |
| 3% | 4.85 |
| 9% | 5.34 |
| 15% | 6.34 |
| 21% | 6.85 |
| 27% | 7.37 |

Table 2. Fit parameters of performance characteristic curves (PCC) and semi-quantitative parameters for both OCPLS models (EUC and W).

| Parameter | EUC Model | W Model |
|----------------------------------|--|--|
| R² | 0.9948 | 0.9937 |
| RMSE (%) | 3.68 | 5.17 |
| Adjusted R-square | 0.9934 | 0.9873 |
| Equation | $\frac{94.98}{(1 + e^{(11.88+(-1.211 \cdot x)})} + 4.67$ | $\frac{-100.2}{(1 + e^{(-3.05+(0.44 \cdot x)})} + 100.9$ |
| CCα (%) | 5 | 2 |
| CCβ (%) | 12 | 12 |
| UR (%) | 5-12 | 2-12 |

Table 3. a) Uncertainty intervals and percentage of samples assigned as adulterated and as uncertain considering two class limits (d_{lower_lim} and d_{upper_lim}). b) Percentage of samples assigned as adulterated considering one class limit ($d_{i,r} \leq 1$) for eucalyptus (EUC) and wild (W) honey.

a)

| Model | Uncertainty interval (d) | % Adulterant assignment | | | | | | |
|-------|--------------------------|-------------------------|----|----|----|-----|-----|-----|
| | | 0% (NA) | 1% | 3% | 9% | 15% | 21% | 27% |
| EUC | 0.69-0.90 | 6 | 0 | 14 | 43 | 100 | 100 | 100 |
| | | 11 | 14 | 0 | 43 | 0 | 0 | 0 |
| W | 0.60-0.90 | 6 | 29 | 14 | 71 | 100 | 100 | 100 |
| | | 37 | 43 | 43 | 14 | 0 | 0 | 0 |

b)

| Model | Uncertainty interval (d) | % Adulterant assignment | | | | | | |
|-------|--------------------------|-------------------------|----|----|----|-----|-----|-----|
| | | 0% (NA) | 1% | 3% | 9% | 15% | 21% | 27% |
| EUC | $d \leq 1$ | 6 | 0 | 0 | 29 | 100 | 100 | 100 |
| W | | 6 | 14 | 14 | 71 | 100 | 100 | 100 |