Garlic adulteration detection using NIR and FTIR spectroscopy and chemometrics

Galvin-King, P., Haughey, S., & Elliott, C. (2021). Garlic adulteration detection using NIR and FTIR spectroscopy and chemometrics. *Journal of Food Composition and Analysis*, 96. <https://doi.org/10.1016/j.jfca.2020.103757>

Published in:

Journal of Food Composition and Analysis

Document Version:

Peer reviewed version

Queen's University Belfast - Research Portal:

[Link to publication record in Queen's University Belfast Research Portal](#)

Publisher rights

Copyright 2021 Elsevier.

This manuscript is distributed under a Creative Commons Attribution-NonCommercial-NoDerivs License

(<https://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits distribution and reproduction for non-commercial purposes, provided the author and source are cited.

General rights

Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Garlic Adulteration Detection using NIR and FTIR Spectroscopy and Chemometrics

Abstract

Garlic (*Allium sativum*) is used daily in a variety of cooking methods worldwide however, it is under threat from economic adulteration. Garlic and possible adulterants such as talc, maltodextrin, corn starch, cornflour, peanut butter powder, sodium caseinate, potato starch, rice flour, cassava and white maize meal were obtained for the development of an adulteration detection method. Near infrared (NIR) and Fourier transform infrared (FTIR) along with chemometrics were used for adulteration detection method development. Principal component analysis (PCA) models were created to establish if there was separation of garlic from the adulterants. Orthogonal partial least squares – discriminant analysis (OPLS-DA) models were then developed resulting in R² and Q² values of 0.985 and 0.914 respectively for NIR. The FTIR values were 0.994 (R²) and 0.964 (Q²). Following validation, the receiver operating curve (ROC) indicated highly accurate models with an area under the curve (AUC) of 0.997 for NIR and 1 for FTIR. The Youden index was calculated at 0.984 and 1 for NIR and FTIR respectively and used to determine the test cut-off value. These results indicate that the NIR and FTIR methods are capable of detecting adulteration in unknown garlic samples and can be used to help protect spice industry from fraud.

Keywords: Garlic; Spice; Economically Motivated Adulteration; Fourier Transform Infrared; Near Infrared; Chemometrics

Declarations of interest

None

25 **1. Introduction**

26 Garlic is one of the world's most popular spices. With over 200 varieties, it is used in a variety of
27 cooking and spice blends (Raghaven, 2007). The European Spice Association (ESA) lists *Allium*
28 *sativum* in the 'ESA list of culinary herbs and spices' (ESA, 2018a). Garlic is from the family
29 Alliaceae (onion family) (Raghaven, 2007) and in dried form, the bulb is the part of the plant used
30 as a spice (Hemphill, 2014).

31 Tridge (<https://www.tridge.com/>) is a global trade ecosystem that provides importers and exporters
32 information on the global market of food and agriculture. It combines professional network and data
33 intelligence for secure cross-border trade on a global scale. According to the market trend data
34 provided by Tridge, a total of 26.57M tonnes of garlic production was reported in 2016. China is the
35 top producer of garlic with 79.8% share in production with 21.2M tonnes reported in 2018. India is
36 in second place at 5.3% of production. China's garlic exports represent 61% of the total exports.
37 Europe imports garlic to a value of US\$423.73 million, and the UK is ranked 8th in the world for
38 garlic imports with a value of US\$57.58 million for 29.23K tonnes reported for 2018 (Tridge).

39 Food fraud is a risk for valuable products such as herbs and spices (Galvin-King et al., 2018). It is
40 therefore important to be aware of preventative measures when dealing in the market. According to
41 the 'Guidance on Authenticity of Herbs and Spices' a number of factors should be considered
42 including product specification, supplier assurance, product type, knowing your supply market, and
43 understanding vulnerabilities in your supply chain (BRC-FDF-SSA, 2016).

44 A vulnerability such as reduction in the crop yield can be an incentive for fraudsters to adulterate
45 garlic. In 2015, poor weather conditions caused stockpiling in warehouses following heavy rain and
46 snow which caused the prices of garlic to spiral upwards (Terazono et al., 2016). In 2010, various
47 reasons caused garlic prices to soar such as a belief that garlic could protect from swine flu, a
48 reduction in the land devoted to growing garlic, and again, poor weather conditions (Branigan,
49 2010). However, when prices remain steady following a decrease in crop output it raises suspicions.

50 It is concerning as it appears that there is more garlic available than has been produced, i.e. when
51 mass balances indicate fraud may be occurring (Brehaut, 2017).

52 Adulteration of garlic is most likely to occur in the powdered form. Garlic powder has the potential
53 to be adulterated with any white powder with talc and chalk powder being identified (Brehaut,
54 2017) as well as starch as adulterants employed (Lee et al., 2014, Lohumi et al., 2015a).

55 Spectroscopy is a technique that can be utilized to detect levels of adulteration that are
56 economically valuable to fraudsters, as observed by the study on the adulteration of oregano (Black
57 et al., 2016) when the level of adulteration ranged mostly from 30-70%. Spectroscopy along with
58 chemometrics is being used increasingly as a powerful tool in the fight against fraud in the herb and
59 spice industry. These techniques allow for rapid analysis, using robust and inexpensive equipment
60 that do not require highly experienced personnel for sample testing once the method is up and
61 running. For the detection of garlic adulteration, Fourier transform infrared (FTIR) was used in the
62 development of a competent method to detect corn starch adulteration in garlic powder with partial
63 least squares regression (PLSR) and variable importance projection (VIP), resulting in an R^2 value
64 of 0.95 and standard error of prediction of 2.56 (Lohumi et al., 2015a). A method for the detection
65 of corn starch in garlic was also developed using Raman with PLSR and VIP with an R^2 value of
66 0.99 and a standard error of correction of 2.16 (Lee et al., 2014). Although a targeted quantitative
67 method has the potential to detect adulteration at lower percentage levels there is also a need for a
68 non-targeted method to detect unknown white powders as well as known adulterants. As garlic can
69 potentially be adulterated with any white powder, a targeted method has its limitations. Although
70 there is a lack of standardisation and legislation for non-targeted techniques, there are guidelines for
71 their validation. Non-targeted methods are valued as important screening techniques (McGrath et
72 al., 2018, US Pharmacopoeia, 2019).

73 In recent years there have been advances in the use of non-targeted methods using spectroscopy and
74 chemometrics for the detection of adulteration in herbs and spices. Examples include the detection

75 of adulteration in oregano (Black et al., 2016), black pepper (Wilde et al., 2019) and saffron
76 (Petrakis and Polissiou, 2017). These methods have the potential to detect any adulterant in the
77 various herbs and spices at an economically viable level.

78 Chemometrics have become a large part of the fraud detection methods along with spectroscopy.
79 Once the spectral data of samples has been collected, chemometrics reduces the
80 multidimensionality of the data and extracts the valuable information from the variables within the
81 data (Kucharska-Ambrożej and Karpinska, 2019). In a recent review of chemometric methods used
82 alongside spectroscopy, it was concluded that the unsupervised technique principal component
83 analysis (PCA) and hierarchical cluster analysis (HCA) were the most commonly used in the
84 detection of adulteration in herbs and spices (Kucharska-Ambrożej and Karpinska, 2019). PCA can
85 be used to view the separation with spectral data information alone, and is often followed by a
86 supervised chemometric model (McGrath et al., 2018). In a study by McGrath et al. (2018) a range
87 of non-targeted food fraud detection techniques and chemometrics used were outlined. It was
88 reported that chemometrics was used in a number of studies for the detection of fraud in a range of
89 products including oils, dairy products, meat, honey and nuts. In this study the chemometrics
90 chosen mostly included the use of the unsupervised technique PCA and the supervised technique
91 orthogonal partial least squares – discriminant analysis (OPLS-DA).

92 The aim of this study was to develop a method for the screening of garlic for economic adulteration
93 using near infrared (NIR) and FTIR. An inexpensive, robust, and rapid method is optimal for the
94 work of detecting economic adulteration in the industry. Both NIR and FTIR benchtop instruments
95 were used to collect the molecular fingerprints of a range of garlic samples and possible adulterants.
96 Chemometrics was used in the development of classification models for the testing method.
97 Following validation of the methods, the criteria for adulteration detection was established based on
98 the chemometric models developed. These methods have the potential to be used by industry to
99 detect the presence or absence of adulteration in new/unknown garlic samples.

100 2. **Materials and Methods**

101 2.1 Sample collection

102 A total of 117 garlic samples were collected from reliable industry sources for the development of
103 an adulteration detection method on two spectroscopic instruments. The samples collected were in
104 various forms including powdered (n=57), minced (n=8), chipped (n=6), granulated (n=40) and
105 flaked (n=6), with some roasted and toasted. The majority of garlic samples came from China,
106 however, India and Egypt were labelled as the origin on one sample each. A garlic (*Allium sativum*)
107 bulb botanical reference material (BRM) (Chromadex, CA, USA) was also obtained for analysis.
108 Twelve white powder adulterants were also obtained; maltodextrin (VWR, Dublin), talc (Sigma-
109 Aldrich, Dorset, UK), corn starch (Sigma-Aldrich, Dorset, UK), rice flour (2), peanut butter
110 powder, sodium caseinate, cornflour, cassava, potato starch and white maize meal were purchased
111 or obtained locally. A corn starch certified reference material (CRM) (National Research Council,
112 Canada) was also included in the sample collection. The chemometric models were built using
113 authentic garlic (n=100) and adulterants (n=12).

114 Three samples of different parts of the garlic plant were also obtained i.e. the root, stem and garlic
115 skins. These samples were analysed to determine if these extraneous parts of the plant would
116 separate from the garlic bulb on a chemometric model following spectral data collection. These
117 models were built using authentic garlic samples (n=100) and a root, stem and garlic skins (n=3).

118 2.2 Sample preparation

119 Once the samples were received, they were milled to a homogenous powder using a planetary ball
120 mill, PM-100 (Retsch, Haan, Germany). All samples, whether, flaked, minced or powdered were
121 treated identically to ensure consistency in the sample format prior to analysis. Approximately 10 g
122 of garlic was added to the grinding jars. The samples were milled for 5 minutes at 500 rpm.
123 Authentic garlic samples (n=117), adulterants (n=12), spiked samples (n=63) and various parts of
124 the plant (n=3) were all prepared in this way and measured in this study.

125 2.3 NIR data acquisition

126 Spectral data collection of garlic and its adulterants was carried out in reflectance mode on the
127 Thermo Antaris II FT-NIR (Thermo Fisher Scientific, Dublin, Ireland). In this analysis,
128 approximately 10 g of each sample was placed into the sample cup (0.5 cm depth) and run on the
129 integrating sphere module of the instrument. Following a background scan, spectra were collected
130 from a rotating sample with a resolution of 8 cm^{-1} in the range of $12000\text{--}4000\text{ cm}^{-1}$. The spectra for
131 each sample were collected in triplicate, and the samples were remixed prior to each collection
132 between replicates. A total of 64 scans were acquired for each spectrum. Prior to chemometric
133 modelling, the three replicates of each sample scanned were averaged.

134 2.4 FTIR data Acquisition

135 Each milled sample was analysed on the benchtop FTIR (Thermo Nicolet iS5, Dublin Ireland) with
136 single bounce iD7 attenuated total reflectance (ATR) accessory and diamond crystal, ZnSe lens and
137 DTGS KBr detector. The experiment set up acquisition parameters also included 0 levels of zero
138 filling, a collection length of 47 seconds, 12415 scan points, N-B strong apodization, mertz phase
139 correction, a laser wavelength of 852 nm as well as a background gain of 4. The spectral data of the
140 garlic and its possible adulterants were collected by placing the milled samples on the diamond
141 crystal sample area of the ATR accessory. Once a sample covered the diamond, a slip clutch
142 pressure tower was placed on top of the sample which secured the sample in place, indicated by a
143 click. This ensured equal pressure was applied to each sample, making the process reproducible. All
144 samples were analysed in triplicate and later averaged. Following a background scan, the spectral
145 data collected ranged from $4000\text{--}550\text{ cm}^{-1}$ with 32 scans at a resolution of 4 cm^{-1} for each analysis.

146 2.5 Chemometrics

147 To extract and correlate chemical information from both the NIR and FTIR data by mathematical
148 and statistical methods, chemometrics was carried out using the software SIMCA 15 (Sartorius,
149 Malmö, Sweden). Pre-processing was initially carried out to reduce the unnecessary variability in

150 the data and focus on the data of interest (Rinnan et al., 2009). To compensate for variation in
151 pathlength, standard normal variate (SNV) was used. First or 2nd derivatives were trialled and the
152 best chosen for each model, and this process was used to reduce baseline shifts and reduce the
153 effects of overlapping bands (Rinnan et al., 2009). Savitzky-Golay (S-G) algorithm was used to
154 smooth the spectra but not to overly reduce the signal-to-noise ratio in the corrected spectra (Rinnan
155 et al., 2009). Pareto scaling was chosen in the development of the models as it reduces intense
156 bands but can emphasize weaker bands that may be of importance. It also maintains the original
157 measurement more than unit variance scaling (Worley and Powers, 2013). Centre scaling on the
158 other hand may cause smaller bands to be masked (Ivosev et al., 2008) by letting the larger bands
159 dominate which, may result in loss of information.

160 In all cases PCA was used to establish if there was a separation of classes with this unsupervised
161 algorithm. The supervised technique OPLS-DA was then chosen to develop the model further.
162 OPLS gives improved model interpretability as it separates the systematic variation in X into two
163 parts, one part that is uncorrelated (orthogonal) to Y, and one part that is correlated (predictive) to
164 Y. OPLS-DA fits an OPLS model using dummy Y variables, one for each class. The dummy
165 variables are a form of numerical variable used in the OPLS-DA models in this study to identify
166 subgroups or class identification. The values are either 1 or 0'. This dummy variable contains a 1 if
167 the sample belongs to the class and a 0 if not (Trygg and Wold, 2002). The PCA models for garlic
168 show each adulterant as its own class. The supervised chemometric models for garlic and its
169 adulterants were binary class models, with 'Garlic' as one class and 'Adulterant' as the other. The
170 class 'Adulterant' included all possible white powder adulterants of garlic. The chemometric
171 models for NIR were based on the spectral range 12000-4000 cm⁻¹. For FTIR the focus was on the
172 range 1800-550, and 4000-2800 cm⁻¹.

173

174 2.6 Validation procedure

175 The validation procedure carried out was based on recommendations for validation of non-targeted
176 methods from the US Pharmacopoeia (2019), Riedl et al. (2015) and McGrath et al. (2018).

177 2.6.1 Internal validation

178 Internal validation was carried out in the chemometric software. The measurement of fit (R²) and
179 the measurement of prediction (Q²) are reported for each model developed. The closer these values
180 are to 1, the better the model. The R² value is based on the explained variation of the real data from
181 the model with the calculation; Residual sum of squares (RSS)= $\sum (\text{observed} - \text{fitted})^2$, $R^2 = 1 - \text{RSS} / \text{SSX}$ where SSX is the overall sum of squares. Internal validation involves cross-validation by
182 removing 1/7th of the data in turn and predicting it against a new model made of 6/7th of the data.
183 This was repeated until all the data was predicted. The predicted residual sum of squares (PRESS)
184 is the squared differences between the observed and predicted values for the data kept out of the
185 model fitting, $\text{PRESS} = \sum (y_{\text{im}} - \hat{y}_{\text{im}})$.
186

187 The overall PRESS/SS, where SS is the residual sum of squares of the previous component is
188 computed in SIMCA. The results of this was represented as a Q² value by dividing PRESS by the
189 initial sum of squares and subtracting from 1. $Q^2 = 1 - \text{PRESS} / \text{SS}$, where SS is the sum of squares of
190 y.

191 2.6.2 External validation

192 A validation set was made up of 100% garlic samples (typical) (n=15) and spiked/adulterated
193 samples (atypical) (n=63). The spiked samples involved two garlic samples with one spiked with
194 corn starch, talc and maltodextrin, and the other spiked with rice flour, peanut butter powder, potato
195 starch and cassava at ten percent intervals from 10-90% (w/w) to make up to 63 spiked samples.
196 Each adulterant was used once for spiking from 10-90% (w/w). All spiked samples were prepared
197 up to 10 g. The 15 100% garlic samples, and the 2 garlic samples used for spiking in the external
198 validation set were not included in the model reference set.

199 The typical and atypical samples were then tested against the binary OPLS-DA model, made up of
200 the classes ‘Garlic’ and ‘Adulterant’. A receiver operating curve (ROC) was developed by plotting
201 the true positive rate (TPR) against the false positive rate (FPR) of the validation set at the various
202 discrimination thresholds (US Pharmacopoeia, 2019). In this study the TPR refers to the number of
203 correctly detected authentic samples and the FPR refers to the number of incorrectly detected
204 authentic samples. The area under the curve (AUC) of the ROC curve illustrates the method
205 performance with the closer the value is to 1, the better, and an AUC of 0.5 indicating a random
206 result (US Pharmacopoeia, 2019). The Youden index (J) was then calculated for the models with
207 the calculation $J = \text{Sensitivity} + \text{Specificity} - 1$ (Greiner and Gardner, 2000) or $J = \text{TPR} - \text{FPR}$ as
208 calculated from the chemometric software. The Youden index ranges from 0-1. The closer these
209 results are to 1, the better (Youden, 1950).

210

211 **3. Results and Discussion**

212 **3.1 Raw data collection**

213 Once the garlic samples and adulterants were milled, and measured on NIR and FTIR spectroscopic
214 instruments, the raw spectral data was collected. Figure 1 (A) and (B) shows the raw spectral data
215 of garlic and four of the possible adulterants (talc, peanut butter powder, rice flour and sodium
216 caseinate) obtained from the NIR (A) and FTIR (B) instruments.

217 In the NIR raw spectral data example, talc is visually the most different with noticeable bands due
218 to OH- stretching vibrations at 7185 cm^{-1} (Petit et al., 2004). According to the herb and spice
219 industry, garlic can be adulterated with any white powder. Therefore, there is the possibility of an
220 inorganic compound such as talc being used. As talc produced a peak at a high wavenumber, the
221 method for the detection of adulteration in garlic includes the full NIR spectral range. The
222 combination C-H stretching observed in the $4500\text{-}4000\text{ cm}^{-1}$, combination N-H and O-H stretching

223 at 5000-4500 cm^{-1} , and C-H and S-H stretching at 6000-5500 cm^{-1} (Aciri et al., 2016) regions
 224 showed great variation between garlic and these adulterants in Figure 1.

225 On the FTIR, talc and sodium caseinate produced the most visually different spectrum. An obvious
 226 band for talc at 3673 cm^{-1} is most likely due to OH- stretching (Petit et al., 2004) and Si-O stretching
 227 at 942 cm^{-1} (Ossman et al., 2014). The bands for sodium caseinate could be due to strong amide I
 228 and II bands from 1700-1500 cm^{-1} (Arrieta et al., 2013). The bands around 1400 cm^{-1} could be
 229 assigned to the carboxylate group (O-C-O) (Arrieta et al., 2013). Garlic also differed from peanut
 230 butter powder, talc and sodium caseinate in the 3000-2800 cm^{-1} region due to C-H stretching
 231 (Lohumi et al., 2015b). In peanut butter powder, the band at 1636 cm^{-1} was in the region where
 232 stretching of C=O and C=C was found (Ossman et al., 2014). A C-O bond was identified at the 995
 233 cm^{-1} region with the highest intensity seen with rice flour (Fan et al., 2012). Both the functional
 234 group region and the fingerprint region were included in the analysis of garlic and its possible
 235 adulterants for the development of chemometric models following FTIR analysis.

236 Following the data collection using both NIR and FTIR, the chemometric models, using an
 237 unsupervised algorithm PCA, was firstly carried out to determine if different classes could be
 238 established between garlic and the adulterants.

239

240 3.2 Chemometric modelling

241 In Figures 2 and 3, the unsupervised PCA chemometric models for garlic and its possible
 242 adulterants are presented. The PCA models here are presented as multi class models, i.e. each
 243 adulterant is separately presented as a class, as well as the garlic class. As PCA is an unsupervised
 244 algorithm, the class information will not affect the model as it creates separation of the samples
 245 based on spectral data alone (McGrath et al., 2018). In Figure 2 (A) and 3 (A), all the adulterants
 246 are included in the model. The first two principal components result in 54.4% and 20.6% variation
 247 in the NIR model and 41.8% and 30.4% variation in the FTIR model respectively. As expected

from the raw data in Figure 1, talc is far removed from the other adulterants and the garlic class, most notably in the NIR spectral data, as is sodium caseinate in the FTIR spectral data. The visual differences in the raw data can be seen in the separation in the PCA chemometric models. For both NIR and FTIR, Figures 2 (B) and 3 (B), once the talc and sodium caseinate are removed from the models, the other adulterants are separated further from the garlic class. Pre-processing for both PCA models included SNV, 1st derivative and S-G (Figure 1 (C) and (D)), with Pareto scaling. As the unsupervised PCA model indicated good separation between the classes, this could then be improved further with the use of a supervised algorithm, with the addition of known class information. The supervised algorithm OPLS-DA was then decided upon to develop a binary classification model. There is clear separation for both NIR and FTIR binary OPLS-DA chemometric models (Figure 4). The classes are composed of 100% garlic samples (Garlic) and all the adulterants combined into one class (Adulterant). Pre-processing for both OPLS-DA models included SNV, 1st derivative and S-G, with Pareto scaling. The measurement of fit (R²) and measurement of prediction (Q²) values for the NIR and FTIR models were calculated as seen in Table 1. These chemometric models were developed and the optimum models were determined based on the R² values and the internal validation carried out represented by the Q² value. These values indicate that the models have good potential for detecting adulteration in unknown garlic samples as the Q² values have results close to 1 for both NIR and FTIR. However, the most accurate process for determining the model's ability to detect adulteration is by carrying out external validation.

268

269 3.3 External validation

External validation began with the randomisation and splitting of data prior to the development of a chemometric model. Therefore, the external validation set, made up of authentic 100% garlic samples (n=15), and samples ranging in adulteration levels at ten percent intervals from 10-90%

273 (w/w) (n=63) could then be run independently against the chosen OPLS-DA models. Once this was
274 carried out, a ROC curve was created. This process was based on the recommendations of the US
275 Pharmacopoeia (2019) and Riedl et al. (2015).

276 The ROC curves for NIR and FTIR as seen in Figure 5 (A) and (B) indicate that chemometric
277 models can detect adulteration in garlic. As indicated by the AUC, both methods gave highly
278 accurate results as determined by Greiner et al. (2000) based on a guideline by Swets et al. (1988)
279 (Table 1). As can be observed in Figure 5 (C) and (D), the predictive scores plot for the external
280 validation set clearly indicates that the greater the level of adulteration, the further the samples are
281 removed from the garlic class. The 10% spiked samples lie closest to the authentic garlic class,
282 while the 90% adulterated samples lie furthest from the authentic class.

283 A Youden index for the NIR and the FTIR methods with chemometrics was derived following
284 external validation (Table 1). This Youden index value was used to determine the cut-off value for
285 the test as the Youden index indicates the maximum differentiating point of the test when false
286 positives and false negatives are of equal value (Greiner et al., 2000). The cut-off value calculated
287 by the Youden index then corresponds to the predicted score value on the classification list
288 produced by the chemometric software. These values were calculated at 0.943 for NIR and 0.982
289 for FTIR. At this cut-off value, the results of the external validation set were classified as
290 adulterated or not i.e. $\geq$ classed as garlic and $<$ classed as adulterated. This could then be used to
291 gain an understanding of the model's capabilities at detecting the various adulteration levels
292 outlined in the external validation set.

293 The results of the external validation had 100% correct classification for authentic garlic detection
294 and adulteration detection in the range of 20-90% for both instruments. Economic adulteration is
295 most probable from 20% upwards for it to be economically worthwhile for the fraudster, with
296 economic adulteration of oregano found to range from 30-70% (Black et al., 2016) and economic
297 adulteration of paprika with spent paprika usually greater than 20% (Bates, 2000). Therefore, the

focus of adulteration detection was from 20% upwards for garlic also. Talc adulteration was not detected at the 10% level using NIR spectroscopy and chemometrics, resulting in 85.7% correct classification at the 10% level for the NIR method. The predicted score value was 0.972, so above the cut-off value for NIR, and therefore classified as garlic. However, all 10% level adulteration was classified correctly with the FTIR method, and therefore overall the FTIR had 100% correct classification for all samples in the validation set. Although the method is built on two classes, garlic and a number of white powder adulterants, any unknown adulterant could potentially be detected if it falls outside the garlic class in the prediction results.

3.4 PCA analysis of extraneous parts of the plant

As there is potential for issues with adulteration of herbs and spices with various parts of the plant, the possibility of detecting extraneous parts of the garlic plant was also investigated. As garlic spice is mainly the bulb of the plant once dried, the root, stem and skins of the garlic were run on the spectroscopy instruments also and an unsupervised PCA chemometric model was created to determine if separation could be established between the classes based on the spectral data (Figure 6). The first two principal components resulted in 50.4% and 23.3% variation in the NIR model and 47.5% and 17.4% variation in the FTIR model respectively. It can be observed that the root and stem are separating from the garlic class more than the skins. At low levels, it is acceptable that extraneous parts of the plant may be incorporated into garlic during production. The ESA sets maximum extraneous levels at 1% (w/w) for spices (ESA, 2018b). If extraneous plant parts are found at high levels, fraud by bulking could be suspected as this may lower the quality of the garlic. The purpose of this PCA model was to determine the potential for the development of a method to detect bulking with extraneous parts of the plant at an economic level as a screening technique. The separation observed in Figure 6 indicated that with addition of more samples and further development, this method could be used in future to ensure high quality garlic.

322

323 4. Conclusion

324 In this study both NIR and FTIR instruments along with a supervised binary classification
325 chemometric model were successful at detecting adulteration. The external validation sets were
326 correctly classified in the 20-90% adulteration range indicating the methods' abilities to be used as
327 screening tools for the detection of economic adulteration. All authentic garlic samples were
328 correctly classified with both techniques. The value of these non-targeted methods ensures that the
329 fraudsters cannot cheat the methods, as any sample that deviates from the authentic class in these
330 methods will be called into question. The future possibilities to advance this work could include the
331 addition of more samples, both authentic and adulterants, to improve robustness and accuracy of
332 the methods. Further possible advances in this field include using similar techniques for various other
333 commodities as well as availing of new advances in technology such as portable and handheld
334 spectroscopy devices for rapid on site analysis. Therefore, these test methods can be used to protect
335 the herb and spice industry as a method of checking supply chains, detecting fraud and deterring
336 fraudsters from criminal activities.

337 Acknowledgements

338 This research was undertaken as part of an industry sponsored PhD studentship, funded by the Herb
339 and Spice Consortium, made up of British Pepper & Spice, McCormick, Bart, Waitrose,
340 Sainsbury's, Morrisons, Asda, M&S, and Tesco.

341

342

343

344

345

346

347 **References**

- 348 Acri, G., Testagrossa, B., Vermiglio, G. (2016). FT-NIR analysis of different garlic cultivars.
349 *Journal of Food Measurement and Characterization*, 10, 1, 127-136.
- 350 Arrieta, M.P., Peltzer, M.A., del Carmen Garrigós, M., Jiménez, A. (2013). Structure and
351 mechanical properties of sodium and calcium caseinate edible active films with carvacrol.
352 *Journal of Food Engineering*, 114, 486-494.
- 353 Bates, L. (2000). Screening Method for the Determination of Defatted Meal in Paprika and Black
354 Pepper. *American Spice Trade Association (ASTA) Standard method ASTA 26.1*.
- 355 Black, C., Haughey, S.A., Chevallier, O.P., Galvin-King, P., Elliott, C.T. (2016). A comprehensive
356 strategy to detect the fraudulent adulteration of herbs: The oregano approach. *Food*
357 *Chemistry*, 210, 551-557.
- 358 Branigan, T. (2010). Garlic bubble leaves bad taste in Chinese mouths. The Guardian. 30th July
359 edn. Retrieved March 28 2020 from:
360 <https://www.theguardian.com/world/2010/jul/30/garlic-chinese-commodity-bubble>.
- 361 BRC-FDF-SSA (2016). *Guidance on authenticity of herbs and spices-Industry Best Practice on*
362 *Assessing and Protecting Culinary Dried Herbs and Spices*. British Retail Consortium, Food
363 and Drink Federation, Seasoning and Spice Association.
- 364 Brehaut, L. (2017). Is there chalk and talcum in your garlic powder? Food fraud expert warns of
365 tainted garlic product., National Post. 18th April, 2017 edn. Retrieved March 27, 2020 from:
366 [https://nationalpost.com/life/food/is-there-chalk-and-talcum-in-your-garlic-powder-food-](https://nationalpost.com/life/food/is-there-chalk-and-talcum-in-your-garlic-powder-food-fraud-expert-is-warning-us-about-tainted-garlic-products)
367 [fraud-expert-is-warning-us-about-tainted-garlic-products](https://nationalpost.com/life/food/is-there-chalk-and-talcum-in-your-garlic-powder-food-fraud-expert-is-warning-us-about-tainted-garlic-products).

368 ESA (2018a). European Spice Association List of Culinary Herbs and Spices, European Spice
 369 Association. Retrieved March 26, 2020 from: [https://www.esa-spices.org/index-](https://www.esa-spices.org/index-esa.html/publications-esa)
 370 [esa.html/publications-esa](https://www.esa-spices.org/index-esa.html/publications-esa)

371 ESA (2018b). European Spice Association Quality Minima Document, European Spice
 372 Association. Retrieved March 26, 2020 from: [https://www.esa-spices.org/index-](https://www.esa-spices.org/index-esa.html/publications-esa)
 373 [esa.html/publications-esa](https://www.esa-spices.org/index-esa.html/publications-esa)

374 Fan, D., Ma, W., Wang, L., Huang, J., Zhao, J., Zhang, H., Chen, W. (2012). Determination of
 375 structural changes in microwaved rice starch using Fourier transform infrared and Raman
 376 spectroscopy. *Starch - Stärke*, 64, 8, 598-606.

377 Galvin-King, P., Haughey, S.A., Elliott, C.T. (2018). Herb and spice fraud; the drivers, challenges
 378 and detection. *Food Control*. 88, 85-97.

379 Greiner, M. and Gardner, I.A. (2000). Epidemiologic issues in the validation of veterinary
 380 diagnostic tests. *Preventive Veterinary Medicine*, 45, 1, 3-22.

381 Greiner, M., Pfeiffer, D., Smith, R.D. (2000). Principles and practical application of the receiver-
 382 operating characteristic analysis for diagnostic tests. *Preventive Veterinary Medicine*. 45, 1,
 383 23-41.

384 Hemphill, I. (2014) Garlic. In *The Spice and Herb Bible* (3rd ed), pp. 292-301. Canada, Robert
 385 Rose Inc.

386 Ivosev, G., Burton, L., Bonner, R. (2008). Dimensionality reduction and visualization in principal
 387 component analysis, *Analytical Chemistry*, 80, 13, 4933-4944.

- 388 Kucharska-Ambrożej, K. and Karpinska, J. (2019). The application of spectroscopic techniques in
389 combination with chemometrics for detection adulteration of some herbs and spices.
390 *Microchemical Journal*. 104278.
- 391 Lee, S., Lohumi, S., Cho, B.K., Kim, M.S., Lee, S.H. (2014). Development of Nondestructive
392 Detection Method for Adulterated Powder Products Using Raman Spectroscopy and Partial
393 Least Squares Regression. *Journal of the Korean Society for Nondestructive Testing*, 34, 4,
394 283-289.
- 395 Lohumi, S., Lee, S., Cho, B. (2015a). Optimal variable selection for Fourier transform infrared
396 spectroscopic analysis of starch-adulterated garlic powder. *Sensors and Actuators B-
397 Chemical*, 216, 622-628.
- 398 Lohumi, S., Lee, S., Lee, H., Cho, B. (2015b). A review of vibrational spectroscopic techniques for
399 the detection of food authenticity and adulteration. *Trends in Food Science & Technology*,
400 46, 1, 85-98.
- 401 McGrath, T.F., Haughey, S.A., Patterson, J., Fauhl-Hassek, C., Donarski, J., Alewijn, M., van Ruth,
402 S., Elliott, C.T. (2018). What are the scientific challenges in moving from targeted to non-
403 targeted methods for food fraud testing and how can they be addressed? – Spectroscopy case
404 study. *Trends in Food Science & Technology*, 76, 38-55.
- 405 Ossman, M.E., Fattah, M.A., Fattah, M.A., Taha, N., Kiros, Y. (2014). Peanut shells and talc
406 powder for removal of hexavalent chromium from aqueous solutions. *Bulgarian Chemical
407 Communications*, 46, 3, 629-639.
- 408 Petit, S., Martin, F., Wiewiora, A., De Parseval, P., Decarreau, A. (2004). Crystal- chemistry of talc:
409 A near infrared (NIR) spectroscopy study. *American Mineralogist*, 89, 319-326.

410 Petrakis, E.A. and Polissiou, M.G. (2017). Assessing saffron (*Crocus sativus* L.) adulteration with
 411 plant-derived adulterants by diffuse reflectance infrared Fourier transform spectroscopy
 412 coupled with chemometrics. *Talanta*, 162, 558-566.

413 Raghaven, S. (2007). *Handbook of Spices, Seasonings, and Flavorings*, 2nd edn, FL, USA: CRC
 414 Press.

415 Riedl, J., Esslinger, S., Fauhl-Hassek, C. (2015). Review of validation and reporting of non-targeted
 416 fingerprinting approaches for food authentication. *Analytica Chimica Acta*, 885,
 417 Supplement C, 17-32.

418 Rinnan, Å, Berg, F.v.d., Engelsen, S.B. (2009). Review of the most common pre-processing
 419 techniques for near-infrared spectra. *TrAC Trends in Analytical Chemistry*, 28, 10, 1201-
 420 1222.

421 Swets, J.A. (1988). Measuring the accuracy of diagnostic systems. *Science*, 240, 4857, 1285-1293.

422 Terazono, E., Li, W., Hornby, L. (2016). *China grapples with garlic price bubble*. Retrieved March
 423 27, 2020 from: <https://www.ft.com/content/flc91d36-9d39-11e6-8324-be63473ce146>

424 Tridge, *Garlic*. Retrieved December 16, 2019 from:
 425 <https://www.tridge.com/intelleigences/garlic/price>

426 Trygg, J. and Wold, S. (2002). Orthogonal projections to latent structures (O-PLS). *Journal of*
 427 *Chemometrics*, 16, 119-128.

428 US Pharmacopoeia (2019). Guidance on Developing and Validating Non-Targeted methods for
 429 Adulteration Detection. *Food Chemicals Codex*, Appendix XVIII, 3S FCC 11.

430 Wilde, A.S., Haughey, S.A., Galvin-King, P., Elliott, C.T. (2019). The feasibility of applying NIR
 431 and FT-IR fingerprinting to detect adulteration in black pepper. *Food Control*, 100, 1-7.

432 Worley, B. and Powers, R. (2013). Multivariate Analysis in Metabolomics. *Current*
433 *Metabolomics*, 1, 1, 92-107.

434 Youden, W.J. (1950). Index for Rating Diagnostic Tests. *Cancer*, 3, 1.

435 Figure 1 Raw spectral data of garlic and some of its possible adulterants on NIR (A) and FTIR (B),
436 and pre-processed data on NIR (C) and FTIR (D).

437

438 Figure 2 NIR PCA chemometric model with garlic and all adulterants, (A) with talc and sodium
439 caseinate excluded (B).

440 t[1]: Principal component 1, t[2]: Principal component 2

441

442 Figure 3 FTIR PCA chemometric model with garlic and all adulterants, (A) with talc and sodium
443 caseinate excluded (B).

444 t[1]: Principal component 1, t[2]: Principal component 2

445

446 Figure 4 OPLS-DA chemometric models for NIR (A) and FTIR (B) analytical techniques

447 t[1]: Principal component 1, to[1]: Orthogonal component 1

448

449 Figure 5 ROC curves and Youden index for NIR (A) and FTIR (B) and a plot of the predictive
450 scores plot for NIR (C) and FTIR (D) following external validation.

451 tPS[1]: Principal component 1 Predicted Score, toPS[1]: Orthogonal component 1 Predicted Score

452

453 Figure 6 PCA chemometric models for garlic and extraneous parts of the plant- root, stem and skins
454 on NIR (A) and FTIR (B)

455 t[1]: Principal component 1, t[2]: Principal component 2

456

457

458

459

460 (A)

461

462 (B)

463

464

465 (C)

466

467

468 (D)

469

470 Figure 1

471

472

473

474

475

476 Figure 2

477

478

479

480

481

482

483

484

485 Figure 3

486

487

488

489

490 (A)

491

492

(B)

493 Figure 4

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

509

510 (A)

(B)

511

512

513 (C)

(D)

515

516

517 Figure 5

518

519

520

521

522 (A) (B)

523

524 Figure 6

525

526 Table 1 The R2, Q2, AUC and Youden index values for the NIR and FTIR methods.

	NIR	FTIR
R2	0.985	0.994
Q2	0.914	0.964
AUC	0.997	1
Youden index	0.984	1

527

528 Table 1

529 Near infrared (NIR), Fourier transform infrared (FTIR), area under the curve (AUC)

530

531

532

533

534

535