LWT - Food Science and Technology 127 (2020) 109368

Contents lists available at ScienceDirect

LWT - Food Science and Technology

journal homepage: www.elsevier.com/locate/lwt

Detection of durum wheat pasta adulteration with common wheat by infrared spectroscopy and chemometrics: A case study

Annalisa De Girolamo[a,*], Marcia Carolina Arroyo[b], Salvatore Cervellieri[a], Marina Cortese[a], Michelangelo Pascale[a], Antonio Francesco Logrieco[a], Vincenzo Lippolis[a]

[a] *Institute of Sciences of Food Production (ISPA), CNR-National Research Council of Italy, Via G. Amendola 122/O, 70126, Bari, Italy*
[b] *INQUISUR (UNS-CONICET), Department of Chemistry, Universidad Nacional del Sur, Av. Alem 1253, B8000CPB, Bahía Blanca, Buenos Aires, Argentina*

ARTICLE INFO

Keywords:
FT-NIR/MIR spectroscopy
Durum wheat pasta adulteration
Rapid method
LDA
PLS-DA

ABSTRACT

Fourier transform (FT) infrared spectroscopy, in combination with Partial-Least Squares Discriminant Analysis (PLS-DA) and Linear Discriminant Analysis (LDA), was used to discriminate commercial durum wheat pasta from Italy and Argentina for common wheat adulteration. Samples were analyzed by both near- and mid-infrared spectroscopy (FT-NIR, FT-MIR) and the performance results were compared. Classification models were developed and validated using Argentinean and Italian durum wheat pasta samples containing common wheat at levels up to 28% and lower than 0.5%, respectively (as determined by ELISA method). The first LDA and PLS-DA models grouped samples into three-classes, i.e. common wheat ≤1%, from 1 to ≤5% and > 5%; while the second LDA and PLS-DA models grouped samples into two-classes using a cut-off of 2% common wheat. The accuracy of the validated models were between 80 and 95% for the three-classes approach and between 91 and 97% for the two-classes approach. In general, the three-classes approach provided better results in the FT-NIR range while the two-classes approach provided comparable results in both spectral ranges.

Results indicate that FT-NIR and FT-MIR spectroscopy, in combination with chemometric models, represent a promising, inexpensive and easy-to-use screening tool to rapidly analyze durum wheat pasta samples for monitoring common wheat adulteration.

1. Introduction

Pasta is one of the most common staple food and a key component of the Mediterranean diet representing an important source of carbohydrates and characterizing a healthy choice among carbohydrate-rich foods for its low glycemic index and high satiating ability. The pasta with a superior quality is that produced from durum wheat (*Triticum durum* Desf) semolina, thanks to the dough excellent rheological properties, cooking quality, and consumer acceptance (Wiseman, 2001). Italy is the Country with both the biggest production (over 3.4 million tons) and consumption (25.3 kg per capita/year) of pasta, while Brazil and Argentina are the biggest pasta producer in South America (1.2 and 0.35 million tons, respectively). Moreover, Argentina is one of the largest consumer of pasta in Latin America (8.3 kg per capita/year) (IPO report, 2014).

According to the Italian law, pasta consumed in Italy is defined as the product obtained exclusively by using durum wheat semolina (DPR n. 187/2001). However, the lower prices of common wheat (*Triticum aestivum* L.) compared to durum wheat, induce some traders to increase benefit economically by the undeclared addition of common wheat flour for durum wheat pasta production. The addition of common wheat flour is an adulteration that leads to a pasta product with a lower resistance to cooking and therefore to a pasta of lower cooking quality (Pauly, Pareyt, Fierens, & Delcour, 2013). As cross contaminations are frequent during growing, harvesting, and flour milling practices, the current Italian law establishes that the maximum amount of common wheat in dry pasta cannot exceed 3%. Furthermore, for export trade, the same Italian legislative decree allows the production of dry pasta with common wheat flour exceeding 3% only if appropriately labeled (DPR n. 187/2001). To ensure a high level of consumer protection in relation to food information, the European Commission (EC) has regulated the information contained on individual food labels to prevent fraudulent practice (EU Regulation N. 1169/2011).

Similarly to Italy, France, Portugal and Greece as well, have adopted a legislation for pasta and have decreed that dry pasta must be produced exclusively (100%) from durum wheat, while other Countries do not adopt any legislation for pasta (UN.A.F·P.A., 2011). Furthermore, in Spain, although pasta can be prepared using durum wheat, common

* Corresponding author.
E-mail address: annalisa.degirolamo@ispa.cnr.it (A. De Girolamo).

https://doi.org/10.1016/j.lwt.2020.109368
Received 14 October 2019; Received in revised form 24 March 2020; Accepted 1 April 2020
Available online 09 April 2020

wheat or their mixtures, when pasta is processed exclusively with semolina of durum wheat it may be classified as of "superior quality".

In Argentina the most appreciated pasta by the consumer is the durum wheat semolina one which consumption increased from 33% in 2014 to 40% in 2017 (UIFRA, 2018). According to the Código Alimentario Argentino (CAA) the manufacture of dry pasta labeled as "semolina pasta" must be made exclusively from *Triticum durum Desf.* semolina (Chapter IX, Art. 708). Furthermore, the complete ingredient list of the packed pasta must be reported on the label (Chapter V, Art. 6.2) (Código Alimentario Argentino (CAA); Ministry of Agriculture, 2012 Nov 2012). This aspect has favored in the last years the gradual consolidation of the pasta production in the Country as well as the reduction of durum wheat pasta import from Italy (Observatory of Economic Complexity (OEC), 2016; UIFRA, 2018).

Information available in literature on food fraud deals with a large number of commodities including cereals and derived product, thus representing a global concern for both safety and economics reasons (Cavanna, Righetti, Elliott, & Suman, 2018; Delwiche, 2016). Therefore, the availability of rapid and reliable methods for the detection of accidental or intentional adulteration of durum wheat pasta with common wheat flour are required by wheat traders, pasta manufacturers and official food control laboratories. Several methods have been proposed to detect the presence of common wheat in durum wheat-based products, including pasta. Some of them are based on the detection of proteins like albumins, gliadin or friabilin by electrophoretic, immunological or chromatographic assays (Knödler, Most, Schieber, & Carle, 2010; Pasqualone, 2011). Another category of methods are based on DNA analysis that utilize the polymerase chain reaction (PCR) or real time PCR to quantify common wheat adulteration (Alary, Serin, Duviau, Joudrier, & Gautier, 2002; Carloni et al., 2017; Casazza et al., 2012; Ibrahim, Al-Hmoud, Al-Rousan, & Hayek, 2011; Pasqualone, 2011).

The development of non-destructive and non-targeted, rapid and easy-to-use methods is being greatly increasing in the last decade. Methods based on infrared spectroscopy (IR) in the near- (NIR) or middle- (MIR) infrared region fulfill these requirements and, in combination with multivariate data analysis, have been largely reported in literature for food quality control and for traceability and authenticity (Cozzolino, 2016; McGrath et al., 2018; Pastor, Acanski, & Vujic 2019; Rodriguez-Saona & Allendorf, 2011). However, only few IR methods were proposed for the detection of durum wheat or pasta adulteration; specifically, Cocchi et al. (2006) described the use of NIR spectroscopy to quantify the degree of adulteration of durum wheat flour with common bread wheat flour, while Kamil, Hussein, Ragab & Khalil (2011) described the use of MIR spectroscopy for identifying and differentiating between wheat varieties and to detect the adulteration of pasta on molecular basis. Furthermore, the use of NIR hyperspectral image technique to detect adulterations of wheat and to discriminate between durum and common wheat kernels products has also been recently reported (Verdú et al., 2017; Vermeulen, Suman, Pierna, & Baeten, 2018; Wilkes et al., 2016). To the best of our knowledge, no applications have been reported to date on the use of IR spectroscopy to detect common wheat in durum wheat pasta. It would be of great value the availability of a rapid method based on IR spectroscopy, in the NIR or MIR range, able to discriminate durum wheat pasta samples based on their common wheat content. Furthermore, considering that the use of FT instrumentation offers several advantages compared to the traditional dispersive IR instruments, the aim of the present paper was to apply for the first time FT-NIR and FT-MIR spectroscopy, together with chemometric analysis, to the detection of durum wheat pasta adulteration with common wheat. The spectral range (i.e. NIR or MIR) giving the best prediction of pasta adulteration was evaluated by comparing two classification models, i.e. partial-least-squares-discriminant analysis (PLS-DA) and linear discriminant analysis (LDA).

2. Materials and methods

2.1. Reagents and apparatus

Ultrapure water was produced by a Millipore Milli-Q system (Millipore, Bedford, MA, USA). ELISA test kits #K381 *Durum EIA* were provided from XEMA-MEDICA Co., Ltd. (Russia). Ethylic alcohol was purchased from Mallinckrodt Baker (Milan, Italy).

2.2. Pasta samples

A total of 280 dry durum wheat pasta samples (500 g each) were purchased from local markets in Italy and Argentina. Specifically, 120 samples of 20 different brands (*i.e. Amato Premium, Barilla, Buitoni, Coop, De Cecco, Divella, Esselunga, Etera, Frediani, Garofalo, Granoro, Granoro dedicato, La Dispensa, La Molisana, Le vie dei Mulini, Pasta Reggia, Riscossa, Selex, Sigma, Tre Mulini*) were produced and marketed in Italy; 154 samples of 16 different brands (*i.e. Bonavita, Cica, Carrefour, Great Value, Knorr, Italia, La Salteña, Lucchetti, Marolio, Matarazzo, Molinos Ala, Molto, Nutregal, Primer Precio, Selection Carrefour, 308*) were produced and marketed in Argentina; 10 samples of different brands (i.e. *Agnesi, Barilla, Buitoni, Colavita, Commendator Giuseppe Ferro, De Cecco, DelVerde, Garofalo, La Majora, Molisana*) were produced in Italy and marketed in Argentina. According to the information reported on the label, Italian and Argentinean pasta samples were made exclusively from durum wheat flour. Each sample was finely ground by the Retsch ZM 200 (Retsch, Haan, Germany) laboratory mill obtaining ground samples with particle size ≤500 μm. Samples were manually homogenized in a bag before FT-IR and ELISA analysis.

2.3. ELISA analysis

Samples were analyzed with the #K381 *Durum EIA* test kit (Xema-Medica Co. Ltd., Russia) according to the procedure provided by the manufacturer (http://xema-medica.com/eng/research/topics/3803.pdf). Two additional quality control durum wheat pasta samples, with expected common wheat levels between 1.1-3.4% and 5.0–9.5%, respectively, were also provided by the manufacturer. Each pasta sample, quality control sample and calibration standard were analyzed by duplicate measurements and the results were expressed as the average value. According to the manufacturer, the limit of quantification (LOQ) of the kit was 0.1% common wheat. The analysis of quality control samples fell within the expected ranges and the coefficient of variation of duplicate analyses was between 1 and 6%.

2.4. FT-NIR spectroscopy analysis

FT-NIR spectra were recorded using the spectrometer Nicolet iS50 FT-IR (Thermo Fisher Scientific Inc., Madison, USA) equipped with an interferometer and an integrating sphere working in diffuse reflection and an indium gallium arsenide detector. Approximately 30 g of ground pasta samples were placed on the rotary sample-cup spinner and spectra were recorded by using 32 interferometer sub-scans in the range between 10,000–4000 cm^{-1}, with a resolution of 8 cm^{-1}.

2.5. FT-MIR spectroscopy analysis

FT-MIR spectra were recorded using the spectrometer Nicolet iS50 FT-IR (Thermo Fisher Scientific Inc., Madison, USA) equipped with an attenuated total reflectance (ATR) accessory consisting of a diamond crystal mounted on a zinc selenide crystal as a focusing element, a pressure applicator ensuring a reproducible pressure during measurements, an interferometer, a potassium bromide beam splitter and a deuterated triglycine sulfate detector. Approximately 50 mg of ground pasta sample was placed on the ATR accessory and four independent measurements, each with a new subsample, were performed. Each measurement was comprised of 16 interferometer sub-scans in the range between 4000–400 cm^{-1}, with a resolution of 4 cm^{-1}. Then, the four spectra were averaged into one spectrum that was pre-treated and used for subsequent multivariate statistics.

2.6. Statistical analysis

The multivariate statistical analysis was conducted with The Unscrambler® X, v10.1 (CAMO Software AS, Oslo, Norway, 2011) for principal component analysis (PCA) and linear discriminant analysis (LDA), and with the Classification Toolbox in Matlab (Mathworks Inc., Natick, Massachusetts, USA) for partial least squares discriminant analysis (PLS-DA) (Ballabio & Todeschini, 2009). Before performing multivariate analysis, spectral data were pre-processed to reduce the spectral baseline shift, improve signal-to-noise ratio and remove light scatter influence. In particular, NIR spectra were pre-processed using mean baseline followed by detrending, while MIR spectra were pre-treated using multiplicative scatter correction followed by detrending (Stuart, 2004).

2.6.1. Principal component analysis (PCA) and linear discriminant analysis (LDA)

The PCA was performed separately on pre-processed FT-NIR and FT-MIR spectra of Argentinean and Italian pasta samples to explore data and to recognize potential clustering (similarities and differences) of the pasta samples as a function of common wheat content in the samples. Eventual outliers were detected by using the graphical tools of the Unscrambler® X software, i.e. the Hotelling T^2 line plot using a critical limit of p-value < 5% and the influence plot, displaying samples with high leverage.

Then, the classification tool LDA was used to classify pasta samples based on their common wheat levels (as determined by the confirmatory ELISA method) by using two different approaches. With the first approach (LDA_1), pasta samples were classified into three groups based on common wheat levels, i.e. class A, common wheat ≤1%; class B (covering the EC maximum permitted level), 1 < common wheat ≤5%; class C, common wheat > 5%. By using this approach samples belonging to class A were identified as acceptable ones, samples of class B should be analyzed by a confirmatory method, while samples of class C were considered rejectable.

With the second approach (LDA_2), pasta samples were classified into two classes using a cut-off of 2% common wheat, i.e. class A, common wheat ≤2%; class B, common wheat > 2%. In this case, all samples belonging to class B should be analyzed by a confirmatory method.

Cut-offs were chosen in order to obtain a balanced number of samples in each class and by assuming a maximum limit of 3% common wheat, like in Italy.

Then, each spectrum was assigned to a class by comparing its Mahalanobis distances from all the classes of the model. The lowest distance between the origin of the discrimination plot and the projection of the sample onto class A, B and C for first approach, or A and B for second approach determined to which class this sample was classified.

2.6.2. Partial least squares discriminant analysis (PLS-DA)

The PLS-DA model was used for pattern recognition of FT-NIR and FT-MIR spectra and is based on the PLS algorithm where the dependent variable *y* is categorical and represents samples class membership (Ballabio & Todeschini, 2009). Also for PLS-DA, two different classification approaches were used. With the first approach (PLS-DA_1), samples with a response value + 1 were attributed to class A (common wheat ≤ 1%); samples with a response value + 2 were attributed to class B (1 < common wheat ≤ 5%) and samples with a response value + 3 were attributed to class C (common wheat > 5%). The number of latent variables (LVs) was chosen as that giving the lowest prediction error in cross-validation (leave-5-out) using the Venetian blinds method. Then, each sample was assigned to the class of the model with the maximum value in the *y* variable.

2.6.3. Evaluation of classification performance

The performance of the classification models developed in the NIR or MIR ranges were evaluated from the results of the analyses of the calibration and validation datasets and expressed in terms of accuracy and sensitivity. All the classification indices were derived from the confusion matrices.

The *accuracy* was defined as the fraction of correctly classified samples with respect to the entire set, and was calculated as follows:

$$Accuracy = \frac{\Sigma CC}{n} x100$$

where *CC* is the number of correctly classified samples and *n* is the number of samples of the entire set.

The *sensitivity* describes the ability of the model to correctly classify samples of the *i-th* class and was defined as:

$$Sensitivity_i = \frac{CC_i}{n_i} x100$$

where CC_i is the number of correctly classified samples of the *i-th* class and n_i is the number of the samples of the *i-th* class.

3. Results and discussion

3.1. Common wheat content by ELISA analysis

Results of ELISA analyses indicated that common wheat was absent in 93% of pasta samples marketed and collected in Italy, while the remaining Italian samples contained common wheat flour between 0.2 and 0.5%, thus indicating that all samples investigated herein fulfilled the requirements of the Italian legislation. Moreover, although the Italian legislative decree allows the export of durum wheat pasta containing common wheat flour at levels higher than 3% only if appropriately labeled (DPR n. 187/2001), only 1/10 Italian pasta samples marketed in Argentina contained common wheat (level of 0.1%). These results were also quite similar to those reported by Casazza et al. (2012) that analyzed Italian pasta samples made from 100% durum wheat using a PCR-based method and showed that the majority of them contained common wheat at levels lower than the tolerated maximum limit of 3%.

Common wheat was found in 112/150 Argentinean pasta samples at level between 0.1 and 28.1%, without reporting it on the label, thus indicating that accidental or intentional adulteration of durum wheat pasta with common wheat occurred (Table 1). Considering that in Argentina the durum wheat pasta should be manufactured exclusively with durum wheat semolina and that any addition of other ingredients should be declared on the label (Código Alimentario Argentino (CAA),), the presence of common wheat flour in pasta samples investigated in

Table 1
Statistical summary of common wheat content (%) in pasta samples from Argentina analyzed by ELISA.

Number of Samples	Range (%)	Mean
38	< LOQ[a]	–
28	0.1–1.0	0.27
25	1.1–2.0	1.69
19	2.1–5.0	3.47
14	5.1–10.0	7.84
30	10.1–28.1	18.26

[a] LOQ, limit of quantification = 0.1% common wheat.

the present study constitutes an adulteration in the market and a fraud when the product would be exported as durum wheat pasta. A similar adulteration was reported by Kelly and Bhave (2007) that applied a PCR-based method to detect common wheat in durum wheat pasta samples marketed in Australia. Results indicated that several samples contained common wheat but it was not listed as ingredient thus suggesting the inaccurate labelling of these products or the cross-contamination at some point in the supply chain.

This is the first time that a survey (about 300 samples) on the adulteration of durum wheat pasta samples from Argentina and Italy is reported.

3.2. Principal component analysis (PCA)

PCA was performed on pre-processed spectra to explore a possible clustering trend between samples. The first 8 PCs described approximately 95% of the total variance in both spectral ranges. By plotting the PC1 vs PC2 sample scores, no visual clustering of the pasta samples based on common wheat content was observed (PC1 and PC2 explained respectively 89% and 5% of the total variance for FT-NIR and 37% and 28% of the total variance for FT-MIR), neither in the score plots of the remaining PCs (data not shown). However, a partial clustering of samples was figured out between pasta samples from Italy and Argentina. Several factors could be responsible for this clustering including the geographical origin and the qualitative characteristics of durum wheat cultivars (such as proteins content, yellow index, ash, gliadin and glutenin content) as well as the production technologies used during pasta manufacturing. These results indicated that PCA can explore spectral data, but it was not able to discriminate these pasta samples based on their common wheat content, therefore, it was necessary to use supervised techniques, i.e. quantitative techniques such as Partial Least Squares (PLS) regression and discriminant techniques, such as LDA and PLS-DA, to build predictive models. Although quantitative approach could be of great interest, the PLS model provided low accuracy results for pasta samples containing low common wheat levels. Based on this we decided to apply qualitative approaches.

3.3. Linear discriminant analysis (LDA)

The first LDA models grouped samples into three-classes, i.e. common wheat $\leq$1%, from 1 to $\leq$5% and > 5% (LDA_1). Considering that all Italian pasta samples contained common wheat at levels < 1%, the first classification models were developed using only Argentinean ones (total of 154 samples). With this approach the Argentinean training set contained 36 samples for class A, 28 samples for class B and 25 samples for class C, while the validation set contained 30 samples for class A, 16 samples for class B and 19 samples for class C (Table 2).

The PCA was applied to compress the information and avoid model over fitting. The number of PCs chosen to get the lowest error in prediction cross validation and then used to build the PC-LDA models was 8.

In the case of FT-NIR range, the LDA_1 model provided sensitivity between 79 and 100% in the training and between 94 and 97% in validation, while the accuracy rate was > 90% in both cases (Table 2). The LDA_1 score plot showed a good separation between the three classes, even though classes B and C showed a higher scattering of samples compared to class A, probably due to the high levels of common wheat in these classes that changed the final composition of the pasta (Fig. 1). Using this approach, 45% of samples would be accepted (Class A), 23% would be analyzed by a confirmatory method (Class B) and 28% would be rejected (Class C).

When the second approach was used to discriminate two classes with a cut-off of 2% common wheat (LDA_2), the training set contained 47 samples for class A and 42 samples for class B while the validation set contained 36 samples for class A and 29 samples for class B. A good accuracy was obtained for both the training (92%) and the validated model (95%), while the sensitivity was between 89 and 95% for the training model and between 94 and 97% for the validated one (Table 3). Using this approach, 43% of samples would be analyzed by a confirmatory method. A clear separation between the two classes was also observed by the LDA_2 score plot (Fig. 2).

In the case of FT-MIR range, the LDA_1 model showed a sensitivity ranging between 68 and 97% in training and between 69 and 87% in validation. The accuracy was of 85% in training and 80% in validation (Table 2). Using this approach, 17% would be analyzed by a confirmatory method. Although these results were less performing compared to FT-NIR, a good separation between the three classes was still observed (Fig. 1).

The LDA_2 model showed a good separation between the two classes as indicated by the accuracy of 94% in both training and validation and the sensitivity that was between 94 and 95% in training and between 92 and 97% in validation (Table 3). Using this model, 43% of samples would be analyzed by a confirmatory method (Table 3). A clear separation between the two classes was also evident in the LDA score plot (Fig. 2).

The better performances obtained by LDA_2 model compared to LDA_1 in the FT-MIR range may be related to the fact that bands from MIR range are caused by fundamental vibrations of lipids, protein, carotenoids, polysaccharides and as a consequence, the MIR fingerprint region is rich in structural information (Rodriguez-Saona & Allendorf, 2011). This kind of discrepancy between the performance of the two models was not observed in the case of FT-NIR range because NIR bands are 10–100 times less intense than their corresponding MIR fundamental bands therefore variability in matrix composition could not have greatly affected model performances.

With the aim of evaluating if the developed models containing only Argentinean samples could also be implemented with Italian ones, additional LDA and PLS-DA classification models were developed by

Table 2
Classification parameters calculated for the training and validation sets of pasta samples from Argentina by applying Principal Component-Linear Discriminant Analysis (LDA_1) and Partial Least Squares-Discriminant analysis (PLS-DA_1) in the FT-NIR and FT-MIR spectral ranges.

Classification models		FT-NIR				FT-MIR			
		Sensitivity[a]			Accuracy	Sensitivity[a]			Accuracy
		Class A	Class B	Class C		Class A	Class B	Class C	
Training	LDA_1	92% (33/36)	79% (22/28)	100% (25/25)	90%	97% (35/36)	68% (19/28)	88% (22/25)	85%
	PLS-DA_1	97% (35/36)	100% (28/28)	92% (23/25)	97%	100% (36/36)	96% (27/28)	88% (22/25)	96%
Validation	LDA_1	97% (29/30)	94% (15/16)	95% (18/19)	95%	87% (26/30)	69% (11/16)	79% (15/19)	80%
	PLS-DA_1	90% (27/30)	93% (13/16)	95% (18/19)	89%	93% (28/30)	63% (10/16)	79% (15/19)	82%

[a] Class A: common wheat ≤ 1%; Class B: 1% < common wheat ≤ 5%; Class C: common wheat > 5%.

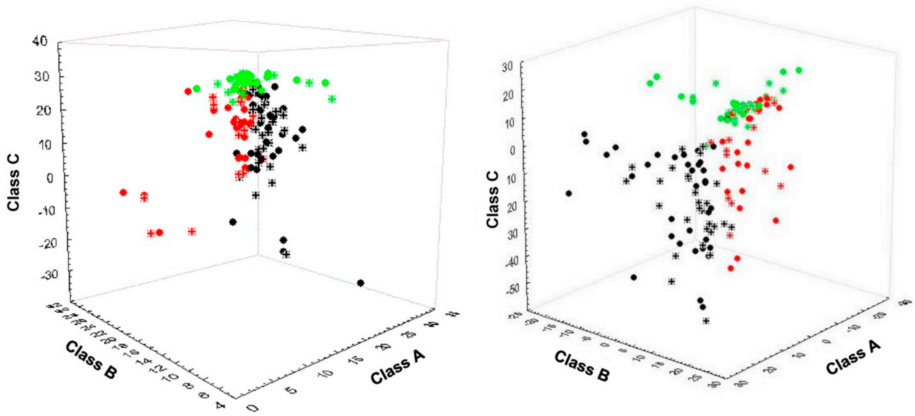

Fig. 1. Predictions results for LDA_1 models developed in the FT-NIR range (left) and FT-MIR range (right) for Argentinean pasta samples. Symbols: circles for training samples; asterisks for validation samples. Colors for each class: black for A, red for B, green for C. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 3
Classification parameters calculated for the training and validation sets of pasta samples from Argentina by applying Principal Component-Linear Discriminant Analysis (LDA_2) and Partial Least Squares-Discriminant analysis (PLS-DA_2) in the FT-NIR and FT-MIR spectral ranges.

Classification models		FT-NIR			FT-MIR		
		Sensitivity[a]		Accuracy	Sensitivity[a]		Accuracy
		Class A	Class B		Class A	Class B	
Training	LDA_2	89% (42/47)	95% (40/42)	92%	94% (44/47)	95% (40/42)	94%
	PLS-DA_2	100% (47/47)	100% (42/42)	100%	92% (43/47)	100% (42/42)	96%
Validation	LDA_2	94% (34/36)	97% (28/29)	95%	92% (33/36)	97% (28/29)	94%
	PLS-DA_2	94% (34/36)	100% (29/29)	97%	92% (33/36)	100% (29/29)	95%

[a] Class A: common wheat ≤ 2%; Class B: common wheat > 2%.

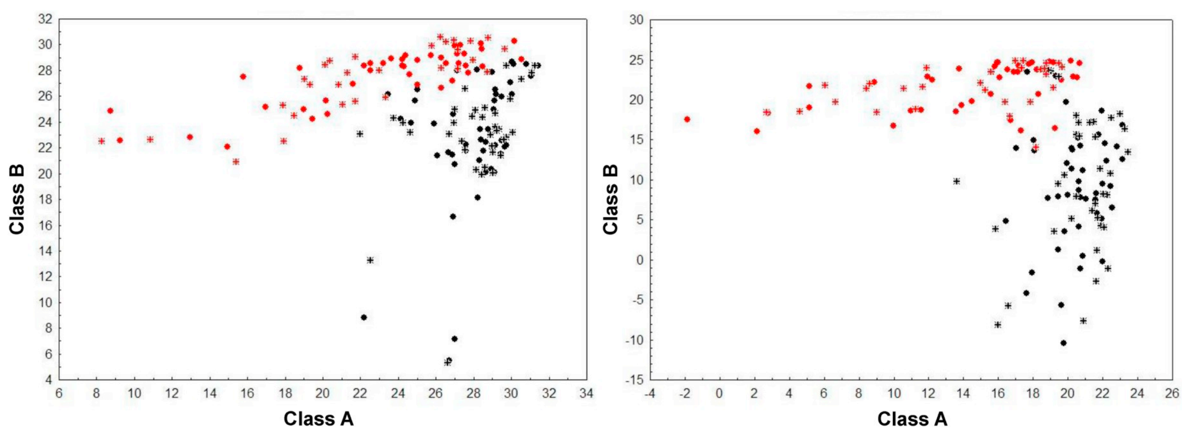

Fig. 2. Predictions results for LDA_2 models developed in the FT-NIR range (left) and FT-MIR range (right) for Argentinean pasta samples. Symbols: circles for training samples; asterisks for validation samples. Colors for each class: black for A, red for B. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 4
Classification parameters calculated for the training and validation sets of pasta samples from Argentina and Italy by applying Principal Component-Linear Discriminant Analysis (LDA_1) and Partial Least Squares-Discriminant analysis (PLS-DA_1) in the FT-NIR and FT-MIR spectral ranges.

Classification models		FT-NIR				FT-MIR			
		Sensitivity[a]			Accuracy	Sensitivity[a]			Accuracy
		Class A	Class B	Class C		Class A	Class B	Class C	
Training	LDA_1	96% (63/66)	68% (19/28)	80% (20/25)	84%	96% (63/66)	82% (23/28)	56% (14/25)	84%
	PLS-DA_1	100% (66/66)	100% (28/28)	100% (25/25)	100%	100% (66/66)	100% (28/28)	100% (25/25)	100%
Validation	LDA_1	98% (59/60)	69% (11/16)	58% (11/19)	85%	90% (54/60)	81% (13/16)	47% (9/19)	80%
	PLS-DA_1	95% (57/60)	69% (11/16)	74% (14/19)	86%	93% (56/60)	88% (14/16)	79% (15/19)	90%

[a] Class A: common wheat ≤ 1%; Class B: 1% < common wheat ≤ 5%; Class C: common wheat > 5%.

Table 5
Classification parameters calculated for the training and validation sets of pasta samples from Argentina and Italy by applying Principal Component-Linear Discriminant Analysis (LDA_2) and Partial Least Squares-Discriminant analysis (PLS-DA_2) in the FT-NIR and FT-MIR spectral ranges.

Classification models		FT-NIR			FT-MIR		
		Sensitivity[a]		Accuracy	Sensitivity[a]		Accuracy
		Class A	Class B		Class A	Class B	
Training	LDA_2	95% (73/77)	76% (32/42)	88%	99% (76/77)	86% (36/42)	94%
	PLS-DA_2	96% (74/77)	93% (39/42)	95%	97% (75/77)	100% (42/42)	98%
Validation	LDA_2	99% (65/66)	79% (23/29)	93%	96% (63/66)	79 (23/29)	91%
	PLS-DA_2	94% (62/66)	86% (25/29)	92%	96% (63/66)	100% (29/29)	97%

[a] Class A: common wheat ≤ 2%; Class B: common wheat > 2%.

including Italian samples in the Argentinean ones. Considering that all Italian samples contained less than 0.5% common wheat, their entire inclusion in the class A would have increased and unbalanced the final number of samples of this class with respect to the other ones, only 30 Italian samples were selected for the training set and other 30 samples for the validation set using the Kennard-Stone (KS) algorithm (Kennard & Stone, 1969). In validation the LDA_1 model showed a sensitivity between 58 and 98% for FT-NIR and between 47 and 90% for FT-MIR and an accuracy of 85% and 80%, respectively (Table 4). Moreover, as observed with the models developed using only pasta samples from Argentina, the two-classes approach using samples from both Argentina and Italy provided better accuracy results between 91 and 93% (Table 5). However, in the case of LDA_1 the number of samples to be analyzed by a confirmatory method for LDA_1 model was 12% for FT-NIR and 14% for FT-MIR spectral range, while for LDA_2 it was 24% for both spectral ranges.

3.3.1. Partial least squares discriminant analysis (PLS-DA)

PLS-DA was applied to test an alternative multivariate statistical approach of classification. The number of latent variables (LVs) guaranteeing the optimal model complexity for PLS-DA_1 was 9 in the FT-NIR range and 12 in the FT-MIR range, while for PLS-DA_2 was 8 in the FT-NIR range and 6 in the FT-MIR range.

In the case of FT-NIR range, the PLS-DA_1 model using only Argentinean pasta samples provided a class sensitivity between 92 and 100% and an accuracy rate of 97% for the training set. In validation, the sensitivity of the model ranged from 90% to 95%, while the overall accuracy was of 89%. With this approach, 20% of samples would be analyzed by a confirmatory method (Table 2, Fig. 3). Good results in terms of sensitivity and accuracy were also obtained for the PLS-DA_2 model; in particular, the accuracy was 100% for the training set and 97% for the validation one (Table 3). Figs. 3 and 4 show the PLS-DA score plots for LDA_1 and PLS-DA_1, respectively.

When PLS-DA_1 model was applied to the FT-MIR spectra the sensitivity was between 63 and 93% while the accuracy was of 82% in validation, with the lowest sensitivity observed for class B, while the number of samples to be analyzed by a confirmatory method accounted to 15% (Table 2, Fig. 3). On the other hands, the PLS-DA_2 model provided better results in terms of sensitivity and accuracy in both training and validation (Table 3, Fig. 4).

Concerning the models including the Italian samples, good validation results were obtained for both PLS-DA_1 (accuracy between 86 and 90%) and for PLS-DA_2 (92–97%) approaches (Tables 4 and 5). Furthermore, in the case of PLS-DA_1 the number of samples to be analyzed by a confirmatory method was 12% for FT-NIR and 15% for FT-MIR spectral range, while for LDA_2 it was 26% for FT-NIR and 31% for FT-MIR spectral range.

Altogether, these data indicate the good ability of FT-IR spectroscopy, in combination with chemometrics, to achieve a good level of discrimination of durum wheat pasta samples either for models using

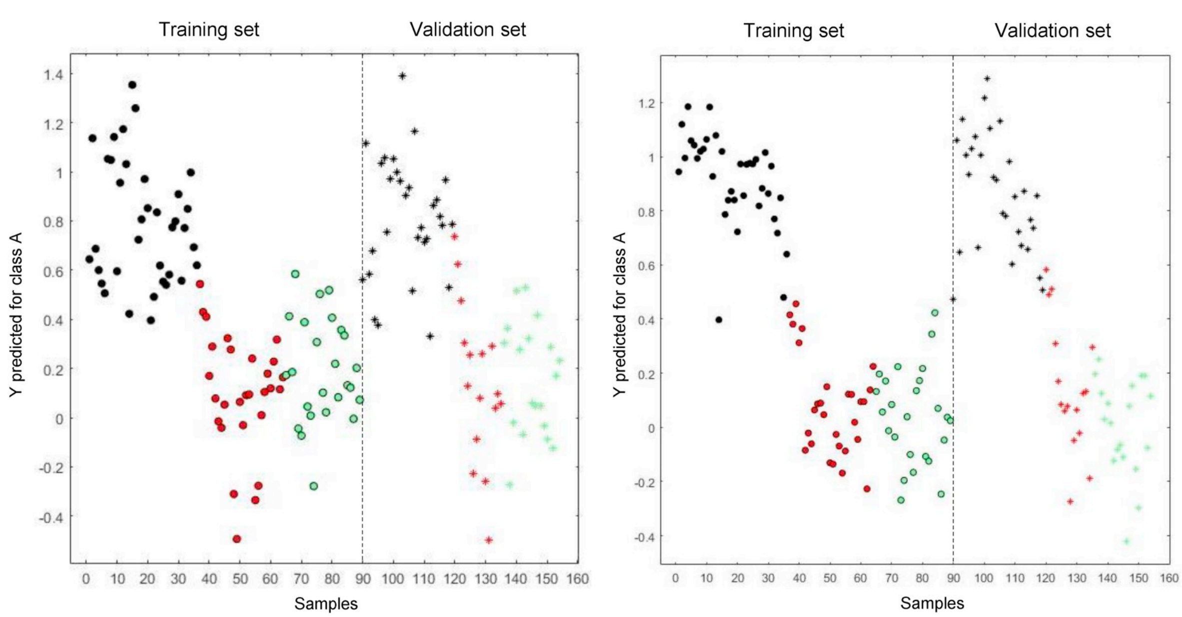

Fig. 3. Predictions results for PLS-DA_1 models developed in the FT-NIR range (left) and FT-MIR range (right) Argentinean pasta samples. Vertical dashed lines indicate the separation between training and validation sets. Symbols: circles for training samples; asterisks for validation samples. Colors for each class: black for A, red for B, green for C. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

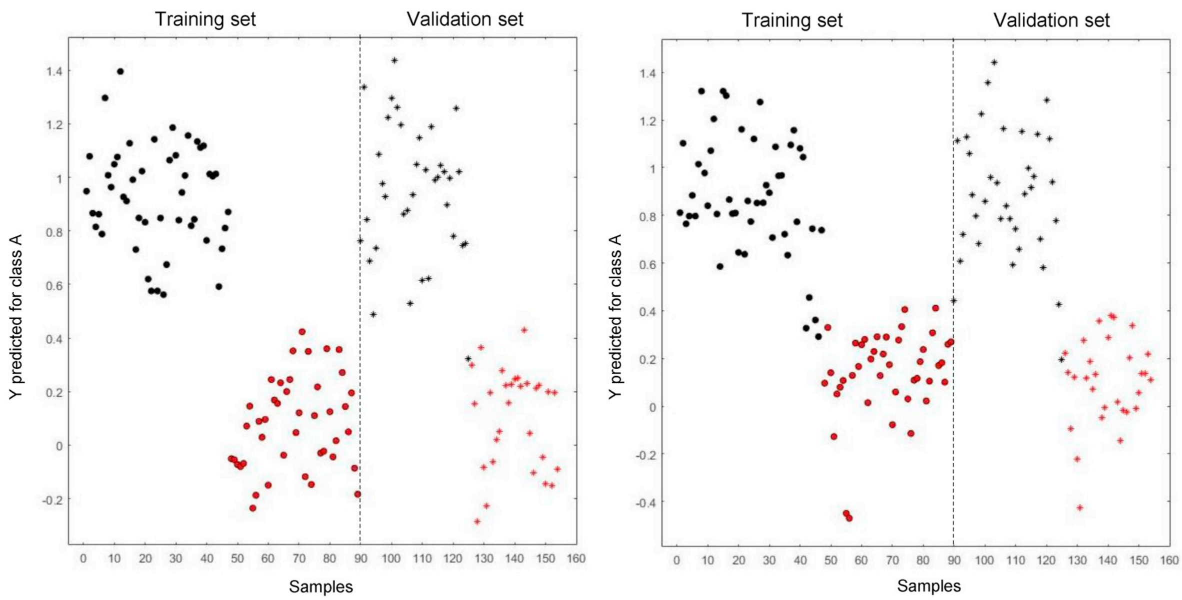

Fig. 4. Predictions results for PLS-DA_2 models developed in the FT-NIR range (left) and FT-MIR range (right) for Argentinean pasta samples. Vertical dashed lines indicate the separation between training and validation sets. Symbols: circles for training samples; asterisks for validation samples. Colors for each class: black for A, red for B. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

only Argentinean pasta samples that for those including Italian ones. Both LDA and PLS-DA models showed comparable performance results. When considering the approaches using three classes for discrimination (i.e. LDA_1, PLS-DA_1), slightly higher accuracy results were obtained for the validated models in the FT-NIR range (85–95%) compared to the FT-MIR range (80–90%). On the other hands, the approaches using two classes for discrimination (i.e. LDA_2, PLS-DA_2) showed similar accuracy results in both ranges (92–97% and 91–97%, respectively).

Findings of the present study were in agreement with those reported by Cocchi et al. (2006) that used NIR spectroscopy to detect adulteration of durum wheat flour with common wheat flour. More recently, Verdú et al. (2017) successfully applied NIR hyperspectral image technique to detect adulteration of wheat flour and bread with cheap grains like sorghum, oats and corn. Results indicated that this technique was able to detect adulterations and a high correlation was observed between wavelengths from specific spectra zones and the physico-chemical properties of samples. Other papers describe the application of NIR spectroscopy combined with PLS-DA to discriminate cocoa powder adulterated with Carob flour or kernels and flours of different wheat species (Quelal-Vasconez, Perez-Esteve, Arnau-Bonachera, Barat, & Talens, 2018; Ziegler et al., 2016). Only one application was reported for application of MIR spectroscopy to the detection of cereal adulteration (Kamil, Hussien, Ragab, & Khalil, 2011).

4. Conclusions

FT-NIR and FT-MIR spectroscopy was used to classify Argentinean and Italian commercial durum wheat pasta for common wheat adulteration by using Partial-Least Squares-Discriminant Analysis and Linear Discriminant Analysis models. Two approaches, classifying samples into three- or two-classes, were used and compared. Performance results indicate that these classification models developed in both spectral ranges could be successfully used to monitor durum wheat pasta adulteration. Furthermore, the approach using three-classes can be considered the more appropriate one by taking into account the reduction of samples to be analyzed by a confirmatory method. Considering that nowadays most of the mills and of the pasta factories routinely use IR techniques for quality and process control applications, it is evident the potential for food industry of using this technique for monitoring both pasta adulteration and quality control.

CRediT authorship contribution statement

Annalisa De Girolamo: Conceptualization, Methodology, Writing - review & editing, Supervision. **Marcia Carolina Arroyo:** Methodology, Formal analysis, Writing - review & editing, Funding acquisition. **Salvatore Cervellieri:** Data curation, Methodology, Formal analysis, Validation, Writing - review & editing. **Marina Cortese:** Conceptualization, Methodology, Formal analysis. **Michelangelo Pascale:** Conceptualization, Methodology, Writing - review & editing. **Antonio Francesco Logrieco:** Conceptualization, Writing - review & editing. **Vincenzo Lippolis:** Conceptualization, Methodology, Writing - review & editing, Supervision.

Declaration of competing interest

The Authors declare that there are no conflict of interests.

Acknowledgements

The authors would like to thank CONICET (Argentina) for the external fellowship for members of the “Research Support Staff Career (CPA)”, XEMA-MEDICA Co. Ltd. (Russia) for kindly providing ELISA test kits, Dr. Christoph von Holst (EC, Joint Research Centre, Belgium) for his statistical contribute and Dr. Erminia Mancini (CNR-ISPA, Italy) for her technical support.

References

Alary, R., Serin, A., Duviau, M. P., Joudrier, P., & Gautier, M. F. (2002). Quantification of common wheat adulteration of durum wheat pasta using real-time quantitative polymerase chain reaction (PCR). *Cereal Chemistry, 79*, 553–558.
Ballabio, D., & Todeschini, R. (2009). Multivariate classification for qualitative analysis. In D. W. Sun (Ed.). *Infrared spectroscopy for food quality analysis and control* (pp. 83–104). Burlington, MA: Academic Press/Elsevier 2009.
Carloni, E., Amagliani, G., Omiccioli, E., Ceppetelli, V., Del Mastro, M., Rotundo, L., et al. (2017). Validation and application of a quantitative real-time PCR assay to detect common wheat adulteration of durum wheat for pasta production. *Food Chemistry, 224*, 86–91.
Casazza, A. P., Morcia, C., Ponzoni, E., Gavazzi, F., Benedettelli, S., & Breviario, D. (2012). A reliable assay for the detection of soft wheat adulteration in Italian pasta is based on the use of new DNA molecular markers capable of discriminating between *Triticum aestivum* and *Triticum durum*. *Journal of Cereal Science, 56*, 733–740.
Cavanna, D., Righetti, L., Elliott, C., & Suman, M. (2018). The challenges from targeted to

non-targeted mass spectrometric methods for food fraud analysis: A proposed validation workflow to bring about a harmonised approach. *Trends in Food Science & Technology, 80*, 223–241.
Cocchi, M., Durante, C., Foca, G., Marchetti, A., Tassi, L., & Ulrici, A. (2006). Durum wheat adulteration detection by NIR spectroscopy multivariate calibration. *Talanta, 68*, 1505–1511.
Código Alimentario Argentino (CAA) Artículos: 643 al 766 - alimentos farináceos - cereales, harinas y derivados.-actualizado al Actualizado al 9/2018. https://www.argentina.gob.ar/anmat/codigoalimentario/, Accessed date: 7 September 2019.
Cozzolino, D. (2016). Near infrared spectroscopy and food authenticity. In M. Espineira, & F. J. Santaclara (Eds.). *Advances in food traceability techniques and technologies: Improving quality throughout the food chain* (pp. 119–136). Cambridge, UK: Woodhead publishing.
Delwiche, S. R. (2016). Advances in the identification of adulterated cereals and cereal products. In G. Downey (Ed.). *Advances in food authenticity testing* (pp. 491–518). Duxford, UK: Woodhead.
DPR n.187. (2001). *Regolamento per la revisione della normativa sulla produzione e commercializzazione di sfarinati e paste alimentari, a norma dell'articolo 50 della legge 22 febbraio 1994.* Italy: Republic President146.
European Commission (2011). Regulation (EU) No 1169/2011 of the European parliament and of the council of 25 october 2011 on the provision of food information to consumers, amending regulations (EC) No 1924/2006 and (EC) No 1925/2006 of the European parliament and of the council, and repealing commission directive 87/250/EEC, council directive 90/496/EEC, commission directive 1999/10/EC, directive 2000/13/EC of the European parliament and of the council, commission directives 2002/67/EC and 2008/5/EC and commission regulation (EC) No 608/2004. *Official Journal, L, 304*, 18–63.
Ibrahim, M. A., Al-Hmoud, N. D., Al-Rousan, H., & Hayek, B. O. (2011). Detection of durum wheat pasta adulteration in the Jordanian market by polymerase chain reaction technology. *American Journal of Food Technology, 6*, 492–499.
International Pasta Organization (IPO) (2014). The world pasta industry status report 2014. http://www.internationalpasta.org/, Accessed date: 7 September 2019.
Kamil, M. M., Hussien, A. M. S., Ragab, G. H., & Khalil, S. K. H. (2011). Detecting adulteration of durum wheat pasta by FT-IR spectroscopy. *Journal of American Science, 7*, 573–578.
Kelly, F., & Bhave, M. (2007). Application of a DNA-based test to detect adulteration of bread wheat in pasta. *Journal of Food Quality, 30*, 237–252.
Kennard, R. W., & Stone, L. A. (1969). Computer aided design of experiments. *Technometrics, 11*, 137–148.
Knödler, M., Most, M., Schieber, A., & Carle, R. (2010). A novel approach to authenticity control of whole grain durum wheat (Triticum durum Desf.) flour and pasta, based on analysis of alkyl resorcinol composition. *Food Chemistry, 118*, 177–181.
McGrath, T. F., Haughey, S. A., Patterson, J., Fauhl-Hassek, C., Donarski, J., Alewijn, M., ... Elliott, C. T. (2018). What are the scientific challenges in moving from targeted to non-targeted methods for food fraud testing and how can they be addressed?-Spectroscopy case study. *Trends in Food Science & Technology, 76*, 38–55.
Ministry of Agriculture. Informe de producto Cadena Pastas Alimenticias. (2012). www.alimentosargentinos.gob.ar, Accessed date: 21 March 2020 Nov 2012.
Observatory of Economic Complexity (OEC) (2016). Which countries import Pasta. https://atlas.media.mit.edu/it/resources/about/, Accessed date: 7 September 2019.
Pasqualone, A. (2011). Authentication of durum wheat-based foods: Classical vs. innovative methods. In M. B. P. P. Oliveira, I. Mafra, & J. S. Amaral (Eds.). *Current topics on food authentication* (pp. 23–39). Kerala (India): Transworld Research Network.
Pastor, K., Ačanski, M., & Vujić, D. (2019). A review of adulteration versus authentication of flour. *Flour and breads and their fortification in health and disease prevention* (pp. 21–35). Academic Press.
Pauly, A., Pareyt, B., Fierens, E., & Delcour, J. A. (2013). Wheat (*Triticum aestivum* L. And *T. turgidum* L. ssp. *durum*) kernel hardness: II. Implications for end-product quality and role of puroindolines therein. *Comprehensive Reviews in Food Science and Food Safety, 12*, 427–438.
Quelal-Vasconez, M. A., Perez-Esteve, E., Arnau-Bonachera, A., Barat, J. M., & Talens, P. (2018). Rapid fraud detection of cocoa powder with carob flour using near infrared spectroscopy. *Food Control, 92*, 183–189.
Rodriguez-Saona, L. E., & Allendorf, M. E. (2011). Use of FTIR for rapid authentication and detection of adulteration of food. *Annual Review of Food Science and Technology, 2*, 467–483.
Stuart, B. H. (2004). Spectral analysis. In John Wiley & Sons Ltd (Vol. Ed.), *Infrared spectroscopy: Fundamentals and applications: Vol. 19*, (pp. 45–70). Chichester, West Sussex PO19 8SQ, England: The Atrium, Southern Gate.
Union of organizations of manufactures of pasta products of the EU (UN.A.F.P.A.). (2001). Pasta legislation in the EU. http://www.pasta-unafpa.org/, Accessed date: 7 September 2019.
Unión de Industriales Fideeros de la República Argentina (UIFRA) (2018). Reporte anual de la Industria fideera Argentina. http://www.uifra.org.ar/, Accessed date: 7 September 2019.
Verdú, S., Vásquez, F., Grau, R., Ivorra, E., Sánchez, A. J., & Barat, J. M. (2017). Detection of adulterations with different grains in wheat products based on the hyperspectral image technique: The specific cases of flour and bread. *Food Control, 62*, 373–380.
Vermeulen, P., Suman, M., Pierna, J. A. F., & Baeten, V. (2018). Discrimination between durum and common wheat kernels using near infrared hyperspectral imaging. *Journal of Cereal Science, 84*, 74–82.
Wilkes, T., Nixon, G., Bushell, C., Waltho, A., Alroichdi, A., & Burns, M. (2016). Feasibility study for applying spectral imaging for wheat grain authenticity testing in pasta. *Food and Nutrition Sciences, 7*, 355–361.
Wiseman, G. (2001). In R. C. Kill, & K. Turnbull (Eds.). *Pasta and semolina technology* (pp. 11–42). Oxford OX2 OEL: Blackwell Science Ltd, Osney Mead.
Ziegler, J. U., Leitenberger, M., Longin, C. F. H., Würschum, T., Carle, R., & Schweiggert, R. M. (2016). Near-infrared reflectance spectroscopy for the rapid discrimination of kernels and flours of different wheat species. *Journal of Food Composition and Analysis, 51*, 30–36.