Taylor & Francis
Taylor & Francis Group

A novel chemometric classification for FTIR spectra of mycotoxin-contaminated maize and peanuts at regulatory limits

Gregor Kos[a], Markus Sieger[b], David McMullin[c], Celine Zahradnik[c], Michael Sulyok[c], Tuba Öner[b], Boris Mizaikoff [b] and Rudolf Krska[c]

[a]Department of Atmospheric and Oceanic Sciences, McGill University, Montreal, QC, Canada; [b]Institute of Analytical and Bioanalytical Chemistry, Ulm University, Ulm, Germany; [c]Department of Agrobiotechnology (IFA-Tulln), University of Natural Resources and Life Sciences, Tulln, Austria

ABSTRACT

The rapid identification of mycotoxins such as deoxynivalenol and aflatoxin B_1 in agricultural commodities is an ongoing concern for food importers and processors. While sophisticated chromatography-based methods are well established for regulatory testing by food safety authorities, few techniques exist to provide a rapid assessment for traders. This study advances the development of a mid-infrared spectroscopic method, recording spectra with little sample preparation. Spectral data were classified using a bootstrap-aggregated (bagged) decision tree method, evaluating the protein and carbohydrate absorption regions of the spectrum. The method was able to classify 79% of 110 maize samples at the European Union regulatory limit for deoxynivalenol of 1750 µg kg^{-1} and, for the first time, 77% of 92 peanut samples at 8 µg kg^{-1} of aflatoxin B_1. A subset model revealed a dependency on variety and type of fungal infection. The employed CRC and SBL maize varieties could be pooled in the model with a reduction of classification accuracy from 90% to 79%. Samples infected with *Fusarium verticillioides* were removed, leaving samples infected with *F. graminearum* and *F. culmorum* in the dataset improving classification accuracy from 73% to 79%. A 500 µg kg^{-1} classification threshold for deoxynivalenol in maize performed even better with 85% accuracy. This is assumed to be due to a larger number of samples around the threshold increasing representativity. Comparison with established principal component analysis classification, which consistently showed overlapping clusters, confirmed the superior performance of bagged decision tree classification.

Introduction

Mycotoxins are secondary metabolites produced by fungal species that infest agricultural commodities such as peanuts, maize and wheat. Globally, mycotoxins are a major driver behind spoiled food with important implications for human and animal health, causing significant financial damage (Wu 2007). The global economic cost due to mycotoxins is estimated to run to billions of US dollars (Swamy et al. 2004). Vardon et al. (2003) estimated an annual range of losses from US$0.5 million to over US$1.5 billion from aflatoxin in maize and peanuts, fumonisins in maize, and deoxynivalenol in wheat. Typically, plants are infected by mycotoxigenic fungi, e.g., during particularly wet growth seasons or during storage, if moisture conditions greater than 14% prevail (Richard 2007) and support fungal growth. Mycotoxigenic fungi can be plant pathogens, such as *Fusarium graminearum*

infecting wheat and maize. *F. verticillioides* and *Aspergillus flavus* will produce mycotoxins on senescent or stressed maize crops. Soil-borne fungi may occur on developing kernels and later produce mycotoxins in storage if conditions permit (Miller 2008). Resulting concentrations in localised 'hot-spots' can be so high that they could render an entire crop batch unsuitable for consumption (Choudhary & Kumari 2010).

Aflatoxins and deoxynivalenol are among the most important toxic fungal metabolites, because of their global occurrence and toxicity. Aflatoxins are mainly produced by *A. flavus, A. parasiticus* and *A. nomius* (Ertas et al. 2011). Aflatoxin B_1 is classified as a Group 1 human carcinogen by the IARC (1993) (Smela et al. 2001) and is frequently detected in peanuts and peanut-based food and feed originating from sub-tropical and tropical regions. However, aflatoxins have recently

been detected in European maize harvested after a summer with similarly hot conditions, which are increasingly linked to climate change (Prandini et al. 2009). Aflatoxins are a cause of human liver cancer and, in high doses, have caused deaths from aflatoxicosis. Significant negative effects of aflatoxin on child growth have been reported, as well as immune modulation (Groopman et al. 1988; IARC 2016).

Deoxynivalenol belongs to the group of B-trichothecenes and is mainly produced by *F. graminearum* and *F. culmorum*. Just like aflatoxins, it is a mycotoxin of global concern, occurring at sometimes high to very high concentrations in the mg kg^{-1} range. Affected commodities are wheat, barley and oats, and maize. Thus it becomes important as a contaminant of food despite its relatively low toxicity. In animals, particularly swine, deoxynivalenol-containing feed starting at levels of 1 μg kg^{-1} as reported by Richard (2007) leads to feed refusal, vomiting and intestinal damage, especially to the kidneys (Pestka 2010).

Because of their global occurrence, legislation for both aflatoxins and deoxynivalenol with other *Fusarium* toxins has been enacted to ensure food and feed chain safety. A summary of European Union (EU) regulatory limits and non-binding guidance values (e.g., for deoxynivalenol in animal feed) regarding aflatoxin B_1 and deoxynivalenol are shown in Table 1. For analytical screening methods to be developed, these limits represent an upper limit regarding classification performance.

Due to increased global trade of agricultural commodities, an expanding number of extreme weather events and climate change-related processes, there is a strong need to increase testing during import, storage

and processing. Current state-of-the-art methodology relies mostly on LC-MS (e.g., Malachová et al. 2014), HPLC-UV and HPLC-FLD (e.g., Rodrigues et al. 2011) for the determination of toxin concentrations. Especially, LC-MS methods are highly sensitive and have been comprehensively validated (e.g., by Malachová et al. 2015) to provide documented accuracy and precision for regulatory purposes. Analyses require advanced laboratory facilities and highly qualified personnel, thus making these methods difficult and costly to use for 'informal' tests by importers, traders, and food- and feed-processing companies. One exception is ELISA (e.g., Turner et al. 2009), which has found application outside regulatory testing, but with ongoing issues regarding cross-reactivities and the need to handle solvents and testing solutions. A review about ELISA and similar assay-based methods was recently published by Chauhan et al. (2016).

Spectroscopy-based rapid screening methods have the potential to fill this gap regarding much-requested 'informal' testing methods. Rather than replacing LC-MS and similar methods, these complementary screening methods provide importers and processors with suitable easy-to-use screening tools without the need for extensive laboratory facilities and highly qualified personnel. Within the EU, costs for sampling and analysis are typically borne by the importer according to Commission Regulation (EC) No. 1152/2009. Therefore, preliminary analyses conducted with rapid screening methods could provide traders with information about potential contamination. By removing contaminated batches right away, the number of samples submitted to regulatory testing can be reduced. As a consequence, rapid methods are suitable to achieve goals of increased testing at, for example, ports of entry and reduce the number of regulatory testing procedures. A large number of simple on-site checks during import, storage and processing, therefore, becomes feasible.

Infrared (IR) spectroscopy-based methods provide simple, non-destructive means to obtain information about a sample. Especially, the near-infrared range (NIR; 13,500 to 4000 cm^{-1}), with its inherent advantages of being insensitive to moisture and transparency of glass containers, greatly simplifies the measurement process. NIR spectroscopy has been successful to measure protein and moisture content as a means to assess grain quality (e.g., by Maertens et al. 2004). While NIR provides overtone frequencies of molecule and functional group vibrations, these are directly accessible using the mid-infrared range (MIR; 4500 to 400 cm^{-1}). Typical MIR absorption frequencies recorded of cereal samples include hydroxyl, methylene, carbonyl (CO) stretching,

Table 1. European Union limits and guidance values (μg kg^{-1}) for aflatoxin B_1 and deoxynivalenol in food and animal feed.

Commodity	Aflatoxin B_1	Deoxynivalenol
Food		
Peanuts (for consumption)	2	
Peanuts (for processing)	8	
Cereals for human consumption	2	750
Cereal-based baby food	0.1	200
Bread and cereal snacks		500
Maize (unprocessed)		1750
Maize (milling fraction < 500 μm)	n.a.	1250
Maize (milling fraction > 500 μm)	n.a.	750
Animal feed		
Feed (except below)	20	8000 (cereals)
Maize by-products		12,000
Calves and lamb feed	10	2000
Dairy feed	5	
Pig feed	20	900

Note: Data are summarised from Commission Regulation (EC) No. 1881/2006, EC Directive 2003/100/EC, amending Directive 2002/32, Commission Recommendation 2006/576/EC and Richard (2007). Data listed for deoxynivalenol animal feed are guidance values.

NH bending and CO stretching vibrations (Kos et al. 2003). The latter (CO and NH vibrations) can further be related to protein and carbohydrates present in the sample. Together with water bands, these spectral features dominate the IR spectrum of foods. Additionally, the fingerprint region of the spectrum between 900 and 400 cm^{-1} provides unique information about the sample, although assignment to common vibrations is possible only in selected cases. Changes due to fungal contamination result in changes due to protein and carbohydrate content, i.e., changes to CO and NH bands, which allow for indirect detection of fungi and their toxins. A summary of state-of-the-art IR spectroscopy-based rapid methods was provided by McMullin et al. (2015). Recent advances in IR source technology with the availability of MIR quantum cascade lasers will significantly boost source power, which is expected to lead to improved signal-to-noise ratios and, therefore, higher sensitivity (Young et al. 2009).

Protein and carbohydrate changes due to fungal contamination result in small changes to the MIR spectrum Kos et al. (2003). Interpretation of the generated IR spectrum is challenging, involving several small changes across the absorption spectrum in the protein and carbohydrate absorption region. Evaluation of single (univariate) absorption band maxima, an approach traditionally used for classification and quantification, is not recommended. Rather, a multivariate strategy, also called 'chemometrics', where one or more spectral windows composed of typically hundreds of wavelengths of overlapping bands are used simultaneously to build a classification or quantitation model, is used (Mark & Workman 2010). While generally considered less intuitive, a wealth of information is provided and small changes to the selected spectral regions can be detected and used in the model. For a general introduction and review to chemometric methods used to assess food quality, see Karoui et al. (2010).

Among the most popular chemometric methods is principal component analysis (PCA), which is used as a tool for data reduction and classification. The method builds a set of new orthogonal variables with maximised variance (loadings) from the original absorbance information and calculates a score for each sample. Score plots starting with those explaining most of the variance reveal patterns and clusters of samples that can be used for classification. Among quantitative methods principal component regression (PCR), a combination of PCA, to remove correlations between variables, and multilinear regression (MLR) are popular as is partial least squares regression (PLS) (Naes et al. 2002; Mark & Workman 2010). KNN is an example of a non-parametric method that counts the number of closely spaced data points and identify its class (e.g., above or below a defined contamination threshold) in a training set and based on a majority vote assigns the class of an unknown (Mariey et al. 2001). Similarly, decision trees, 'flow-chart'-like models, can be built with each node testing a property of the data (Quinlan 1986). While these methods tend to create overfitted models (i.e., too closely matching the training set and leading to model instability), bootstrap aggregation (also known as 'bagging') combines data from randomly generated training trees. Variance from noise and error-prone data is minimised, providing a more stable classifier (Breiman 1996).

Multivariate methods have been used extensively to assess food quality and identify and quantify major constituents of foods. Examples are the assessment of protein content in wheat (Williams & Sobering 1993), the determination of fat in milk samples (Inon et al. 2004), the pork content in chicken meat (Al-Jowder et al., 1997) and the classification of modified starches (Fernandez Pierna et al., 2005). Microorganisms such as lactic acid bacteria have also been identified using MIR spectroscopy (Amiel et al. 2000).

Among minor and trace constituents, MIR spectroscopy with chemometric data analysis plays an increasing role as an effective tool to screen for the presence and quantities of contaminants. Table 2 provides an overview of reports published using MIR methods and chemometrics for the determination of trace components in food and feed products.

The application of NIR spectroscopic methods dominates reports in the scientific literature, because of widespread use due to simple instrumentation. MIR-based techniques follow with the occasional use of Raman spectroscopy and photo-acoustic spectroscopy, the latter of which is quite suitable for single-kernel analysis.

Previous work was primarily focused on developmental aspects, often with relatively few samples in the dataset or large mycotoxin concentration differences in order to demonstrate the feasibility of the approach (Kos et al. 2003). Large concentrations also facilitated the initial identification of changes to the IR spectrum, related to changes in protein and carbohydrate concentration due to fungal infection. For application of these methods in real-life environments, it is necessary to have a large number of samples at relevant concentrations (e.g., at the regulatory limit) to reflect accurately the inherent variability of the dataset. To our knowledge, the detection of aflatoxins in peanuts at legal limits has not yet been attempted. Similarly, reports for classification of DON-contaminated maize in the

range of EU legal limits are not yet available. Furthermore, Table 2 shows that PCA and PLS techniques are highly popular for classification and quantification, respectively, compared with techniques such as KNN and decision trees. The latter have not yet been investigated and tested extensively.

In this study we present classification results for deoxynivalenol in maize and, for the first time at trace levels, aflatoxin B$_1$ in peanuts using MIR-ATR spectroscopy with chemometrics. Specifically, we compared results using PCA and, also a first, a bagged decision tree methodology for classification. Reference data were obtained employing a single, well-documented LC-MS method. Classification limits for both mycotoxins follow EU regulations, and a large number of samples were investigated at relevant concentration ranges to allow for full validation of the multivariate approach.

Materials and methods

Sampling and sample preparation

Maize samples were provided by Saatbau Linz (Linz, Austria; SBL) and the Cereal Research Centre (Szeged, Hungary; CRC). All samples were maize hybrids with dent or flint kernel types used for kernel production in the field. Either, natural infection, generally yielding low mycotoxin contaminations, or artificial infection by silk injection or toothpick inoculation for high mycotoxin concentrations, with *F. graminearum, F. verticillioides* or *F. culmorum* was performed. All injection and inoculation work was done in the field before harvest. Silk channel injection was performed 5 days after 50% silking. Maize ears were injected with 2 ml

Fusarium suspension at a concentration of 5×10^5 conidia ml^{-1} in the silk channel. *Fusarium* conidial suspension was produced with the bubble breeding method in mung bean broth (Mesterházy 1977). Small aliquots were frozen at –80°C and quickly thawed at 35°C before use. Inoculation with the toothpick method was performed 10 days after 50% silking. A hole was drilled in the centre of the ear, and a tooth pick overgrown with *Fusarium* isolate was inserted (Chungu et al. 1996). Harvested samples were milled (Romer, Union, MO, USA) and spectral analysis was carried out on the 100–250 μm sieve fraction (Retsch, Hahn, Germany).

For peanuts, 92 different, naturally infected samples from public markets in three different African countries (Tanzania, Mozambique and Burkina Faso) were collected. Each sample was analysed individually and no mixing was performed. Peanuts were manually ground to an oily paste with a mortar and pestle prior to analysis.

Infrared data acquisition

MIR spectra were generated with a portable Fourier transform IR spectrometer equipped with an attenuated total reflection (FTIR-ATR) unit (Bruker Alpha, Karlsruhe, Germany). A diamond crystal with one internal reflection served as an ATR element (Bruker, Platinum ATR). For improved portability an L-alanine doped triglycine sulfate (DLaTGS) detector was employed. Spectra were recorded in the MIR range (4000 to 575 cm^{-1} for maize; 475 cm^{-1} as the lower limit for peanuts) with a resolution of 4 cm^{-1}. Spectra

Table 2. Infrared spectroscopic methods and chemometrics employed for the analysis of minor constituents in foods.

Commodity	Parameter	Instrument	Chemometrics	Reference
Sugar	Polysaccharides	MIR	PCA	Cerna et al. (2003)
Maize	Deoxynivalenol	MIR-ATR	PCA	Kos et al. (2003)
Legumes	Minerals	NIR-reflectance	MPLS	Cozzolino and Moron (2004)
Wine	Phenolic compounds	NIR	PLS	Cozzolino et al. (2004)
Wheat	Deoxynivalenol	MIR-ATR, DRIFT	PLS, MLR	Abramovic et al. (2007)
Milk	Melamin	NIR, MIR-ATR, DRIFT	PLS	Mauer et al. (2009)
Maize	Aflatoxin B$_1$	NIR-reflectance	PLS	Fernandez-Ibanez et al. (2009)
Wheat	Deoxynivalenol	NIR	PLS	Firrao et al. (2010)
Maize	Fumonisins	MIR	ANN	Draganova et al. (2010)
Maize	*Fusarium* spp.	NIR-DRIFT	SIMCA, ANN	Vermeulen et al. (2012)
Wheat	Ergot bodies	NIR	PLSDA, SVM	Williams et al. (2012)
Maize	*F. verticillioides*	NIR	PCA	Sirisomboon et al. (2013)
Rice	Aflatoxigenic fungi	NIR	PLS	De Girolamo et al. (2014)
Wheat	Deoxynivalnol	NIR	LDA	Qiang et al. (2014)
Rice	Aflatoxin B$_1$	NIR	PLS	Lee et al. (2015)
Maize	Aflatoxins	NIR, MIR, Raman	MLR, PCR, PLS	Lee and Herrman (2015)
Maize	Fumonisins	SERS	KNN, LDA, PLS	Kaya-Celiker et al. (2016)
Peanut seeds	*Aspergillus* spp.	PAS	LDA	

Note: ANN, artificial neural networks; ATR, attenuated total reflection; DRIFT, diffuse reflectance Fourier transform infrared spectroscopy; KNN, K-nearest neighbour classification; LDA, linear discriminant analysis; MIR, mid-infrared spectroscopy; MLR, multilinear regression; NIR, near-infrared spectroscopy; PAS, photo-acoustic spectroscopy; PCA, principal component analysis; (M)PLS, (modified) partial least squares regression; PLSDA, partial least squares discriminant analysis; SERS, surface-enhanced Raman spectroscopy; SIMCA, soft independent modelling of class analogy; SVM, support vector machines.

AQ16

were collected from 64 scans with a sample and reference gain of 1. Phase correction was applied using a power spectrum with a phase resolution of 32 and Blackman-Harris 3-Term apodisation function. Solvents employed for cleaning were iso-propanol for maize and methanol, acetone and iso-propanol (applied sequentially) for peanuts.

Typically, 10 repeat measurements, each with a new subsample, were performed on the ground and sieved maize and ground peanut samples. While a spectral window from 1800 to 800 cm^{-1} was employed for chemometric modelling, the acquisition of the full MIR range was feasible by using an FTIR instrument, maximising spectral information content for initial exploratory data analysis for the identification of suitable spectral windows. See Figures 1(a) and (b) for typical repeat spectral measurements without pretreatment. Carbohydrate and protein-related vibrations at 1000 and 1500 cm^{-1}, respectively, are prominent spectral features. Upon fungal infestation, these bands are subject to changes due to changes in protein and carbohydrate content (Boyacioğlu & Hettiarachchy 1995; Kos et al. 2003; McMullin et al. 2015). For the purpose of chemometric modelling, these changes were then
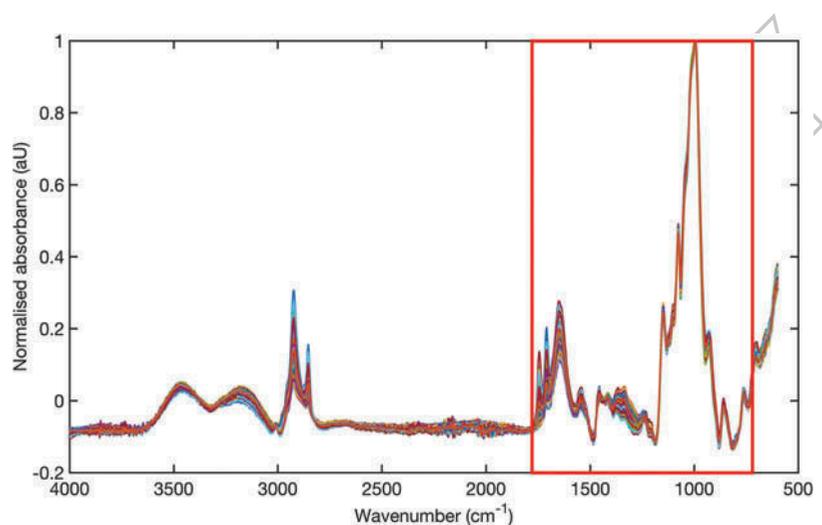

Figure 1(a). MIR spectra of maize; repeat measurements of 10 different subsamples. The (boxed) spectral window between 1800 and 800 cm^{-1} was used for chemometric modelling.
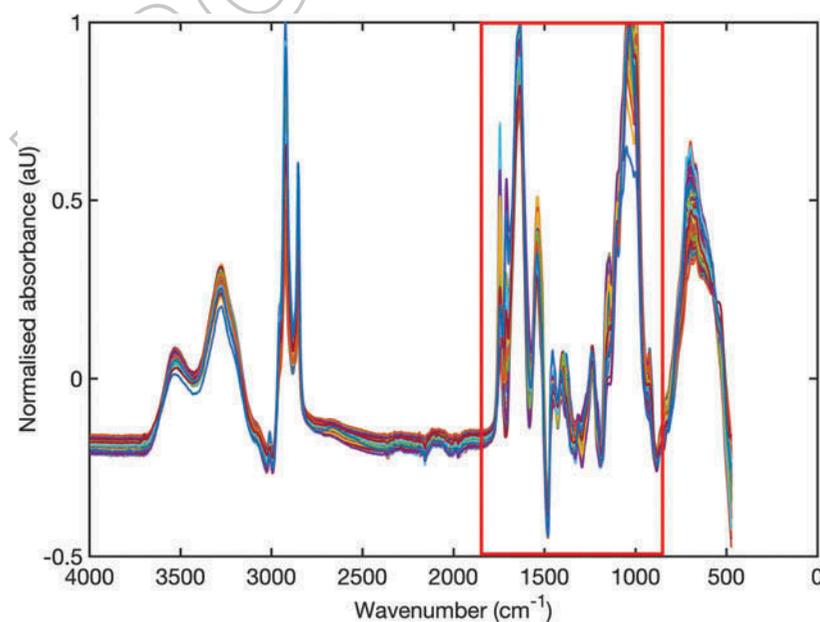

Figure 1(b). MIR spectra of peanuts; repeat measurements of 10 different subsamples. The (boxed) spectral window between 1800 and 800 cm^{-1} was used for chemometric modelling.

related to toxin concentrations, since these are the regulated parameters and facilitate practical application of the developed model.

345 The less specific CH and OH-related bands at approximately 2900 and 3400 cm^{-1} are also clearly visible, but were not used for data analysis for their lack of specificity (Kos et al. 2003). Band assignments for spectra are available in Table 3.

LC-MS/MS reference measurements

350 Homogenised samples of 5 g were extracted with 20 ml of extraction solvent (acetonitrile/water/acetic acid 79:20:1, v/v/v) for 90 min on a rotary shaker (GFL 3017, Burgwedel, Germany). After a 1 + 1 (v/v) dilu-
355 tion using dilution solvent (acetonitrile/water/acetic acid 79:20:1, v/v/v), the diluted extracts were injected without further pretreatment.

The chromatographic method, chromatographic and mass spectrometric parameters are described by
360 Malachová et al. (2014). Briefly, LC-MS/MS screening of target microbial metabolites was performed with a QTrap 5500 LC-MS/MS System (Applied Biosystems, Foster City, CA, USA) equipped with TurboIonSpray electrospray ionisation (ESI) source and a 1290 Series
365 HPLC System (Agilent, Waldbronn, Germany). Chromatographic separation was performed at 25°C on a Gemini C-18-column, 150 × 4.6 mm i.d., 5 μm particle size, equipped with a C18 4 × 3 mm i.d. security guard cartridge (Phenomenex, Torrance,
370 CA, USA).

ESI-MS/MS was performed in the time-scheduled MRM mode both in positive and negative polarities in two separate chromatographic runs per sample by scanning two fragmentation reactions per analyte. The
375 MRM detection window of each analyte was set to its expected retention time ±23 and ±29 s in positive and negative modes, respectively. Confirmation of positive analyte identification was obtained by the acquisition of two MRMs per analyte (with the exception of

380 moniliformin which exhibited only one fragment ion). This yielded 4.0 identification points according to European Commission Decision 2002/657 (European Commission 2002). In addition, the LC retention time and the intensity ratio of the two
385 MRM transitions agreed with the related values of an authentic standard within 0.1 min and 30% rel., respectively. The accuracy of the method is verified on a routine basis by regular participation in proficiency testing schemes (Malachová et al. 2015). The LODs
390 were 1.2 μg kg^{-1} for deoxynivalenol and 0.24 μg kg^{-1} for aflatoxin B$_1$.

Infrared data classification and validation

These acquired IR and LC-MS/MS reference data formed the base for the modelling exercise presented here. Classification thresholds of 1750 and 500 μg kg^{-1}
395 were chosen for deoxynivalenol in maize and 8 μg kg^{-1} for aflatoxin B$_1$ in peanuts in agreement with European Union legislation. Chemometric modelling and plotting was carried out using Matlab (version 2015a),
400 including data import, pre-processing, spectral window selection, chemometric modelling and creation of spectral plots. Due to the heterogeneity of the maize samples provided with regard to multiple variables (e.g. variety, fungal infection, etc.) (Table 4) statistical eva-
405 luation was conducted using several subsets. Also, different varieties of the same commodity have different protein and carbohydrate contents in the original, non-contaminated, state. In order to test the influence of these parameters on the robustness of the developed
410 models, a sub-setting routine was implemented to choose specific combinations for testing. For example, only maize samples of the variety CRC infected with *F. graminearum* were chosen for detailed analysis before expanding the model by including the second SBL
415 variety. Similarly, other fungal species were removed and added to test model robustness. For peanuts, less sample information was available and the dataset was subset by country.

Data pretreatment consisted of baseline correction,
420 averaging, mean-centring, normalisation and calculation of the first derivative. A bagged decision tree classification of pretreated spectra yielded the best classification results for the chosen datasets and classification thresholds. PCA was carried out as a tool for
425 observation of cluster formation, thus serving as a means for comparison, since PCA represents a well-established method in the scientific literature (see Table 2 for recent studies).

The bagged decision tree model was developed using
430 the Statistics and Machine Learning Toolbox for Matlab

Table 3. Band assignments for MIR vibrations in maize and peanuts.

Vibration	Wave number (cm^{-1})
–CH *trans* stretching	3025
–CH *cis* stretching	3006
CH (-CH$_2$-) asymmetric stretch	2924
CH (-CH$_2$-) symmetric stretch	2854
C–O ester stretching	1745
C–O fatty acids	1710
Amide I	1650
Amide II	1540
-C-H stretching of CH$_2$ and CH$_3$, fatty acids	1490–1425
-C-H symmetric bending methyl groups	1410–1330
Ring vibrations of carbohydrates	1030

Table 4. Summary of sample information for maize and peanuts.

Commodity	n	Toxin	Concentration range (mg kg^{-1})	Average (mg kg^{-1})	SD (mg kg^{-1})	Fungi	Infection method
Maize, n = 184							
CRC	90	DON	147–50,160	8102	12,503	*F. gram.*	Natural
						F. vert.	Silk inj.
						F. cul.	Toothpick
SBL	94	DON	< LOD–43,520	10,750	14,850	*F. gram.*	Natural
						F. vert.	Silk inj.
						F. cul.	Toothpick
Peanuts, n = 92							
Burkina Faso	35	AFB1	< LOD–7960	615	1772	n.a.	Natural
Mozambique	47	AFB1	< LOD–10,624	708	1678	n.a.	Natural
Tanzania	10	AFB1	< 10–98	44	28	n.a.	Natural

Note: DON, deoxynivalenol; AFB1, aflatoxin B$_1$; *F. gram., F. graminearum; F. vert., F. verticillioides; F. cul., Fusarium culmorum*. Infections methods were 'natural'; silk injection ('silk inj.') and by toothpick inoculation ('toothpick'). LOD was determined to be 1.5 µg kg^{-1} for DON and 0.24 µg kg^{-1} for AFB1 from Malachová et al. (2014).

'fitensemble' routine (version R2015a) with 300 ensemble learning cycles. For each commodity, the trained classifier was then subjected to a cross-validation with 10 folds. Classification accuracy was calculated as: (1 – classification loss) × 100, thus reported as classification accuracy (%). Classification loss was calculated as the loss averaged over all folds during cross-validation.

The use of cross-validation aimed at using the full sample variability provided by the dataset and conducting multiple validation runs with training and testing subsets to calculated classification accuracy. Thus we avoided a major drawback inherent with training and test validation (also called the 'split-sample approach'), which would have resulted in relatively few samples, especially in the concentration range around the classification threshold. A consequence would have been a lack of representativity in the training and test sets, respectively. Cross-validation has been described as an 'extension of the split-sample method' (Steyerberg et al. 2001) since several random subsamples are selected as training sets for model development and then tested on the test sets, i.e., the remaining data (Refaeilzadeh et al. 2009). Also, in this study the training routine was performed at least five times to investigate further model stability. Since the bagged decision tree method always starts with different starting conditions through bootstrap aggregation (Breiman 1996), results are expected to be slightly different. However, classification accuracy differences between runs were always less than 0.6%, thus serving as a further indicator of model stability.

Results and discussion

We have investigated the classification ability of a decision tree at 1750 µg kg^{-1} for deoxynivalenol in maize and 8 µg kg^{-1} for aflatoxin B$_1$ in peanuts. These corresponded to regulatory limits set by the EU for unprocessed maize and peanuts destined for processing (Table 1). The limit for peanuts also includes shelled peanuts used here for FTIR measurements. For determination of the aflatoxin content, all the contamination is assumed to be on the edible part, as stated by Commission Regulation (EC) No. 165/2010. In order to study further the performance of the decision tree classifier, we also tested the data with a 500 µg kg^{-1} threshold for deoxynivalenol in maize. For comparison, corresponding PCA results were calculated.

Figures 2(a) and (b) provide the confusion matrix for decision tree classification results for maize. Among a total of 110 samples in the dataset, 17 samples with deoxynivalenol concentrations < 1750 µg kg^{-1} (first line) were correctly classified and 17 samples incorrectly classified. For samples with deoxynivalenol concentrations > 1750 µg kg^{-1} (second line), six samples were incorrectly classified and 70 samples correctly assigned as contaminated. Overall classification accuracy after cross-validation was 79%, i.e., 87 out of 110 samples were correctly classified for the 1750 µg kg^{-1} classifier. For the 500 µg kg^{-1} threshold in Figure 2(b), the results changed to three correct classifications and 15 incorrectly classified samples for concentrations < 500 µg kg^{-1}. Only one sample was incorrectly classified as above the threshold, while 91 samples with concentrations > 500 µg kg^{-1} were correctly classified. In this case, overall classification accuracy was 85% after cross-validation (94 out of 110 samples). Score plots of principal components 1 and 2 show significant overlap between tentatively assigned clusters of samples with contamination above and below the threshold of 1750 µg kg^{-1} (Figure 3(a)). In an analogous fashion, Figures 3(b) and (c) provides decision tree and PCA results for peanut samples. Note that despite the much lower classification limit of 8 µg kg^{-1}, model performance was in a similar range as maize with a classification accuracy of 77% for 92 peanut samples (Figure 3(b), 71 samples correctly classified). Again, a tentative cluster assignment for PCA calculations showed significant
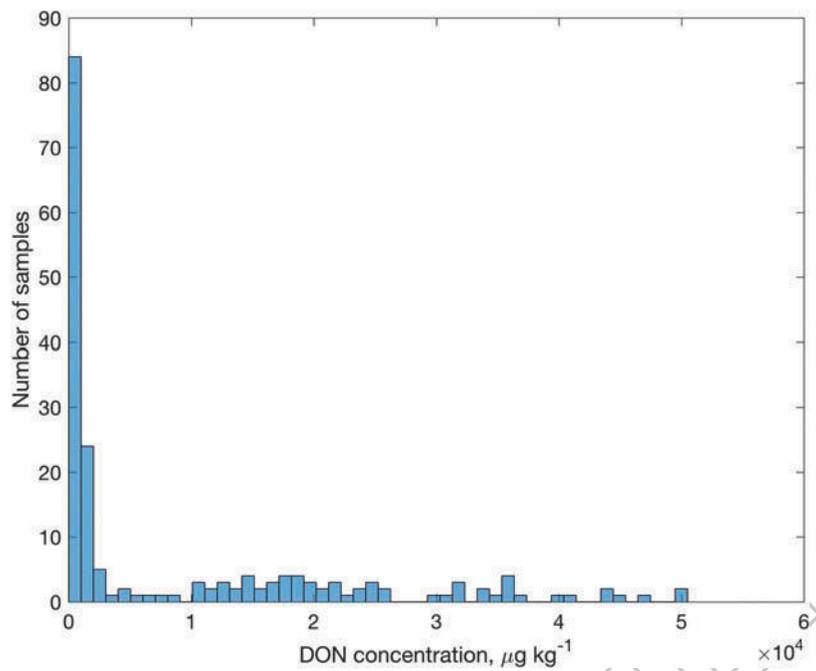
Figure 2(a). Histogram of (a) DON concentrations in maize samples investigated. 823×617mm (72 × 72 DPI)
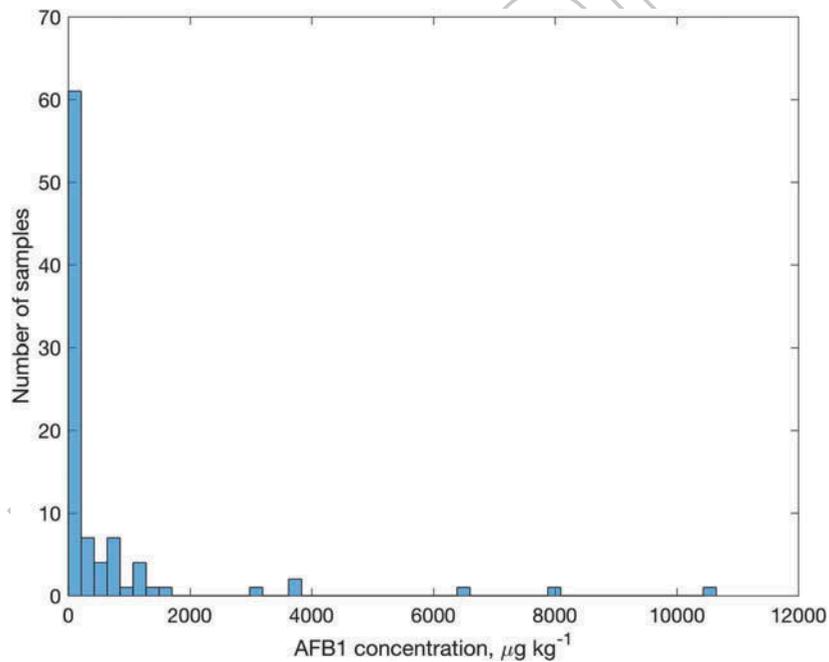

Figure 2(b). Histogram of (b) AFB1 concentrations in peanut samples investigated. 823×617mm (72 × 72 DPI)

overlap between clusters of samples with contamination above and below the threshold of 8 μg kg^{-1} (Figure 3(c)). To our knowledge this is the first report of the classification of MIR spectra at these low concentrations using a decision tree method.

Both concentration limits for maize and peanuts used for classification above and below the chosen thresholds of 1750 μg kg^{-1} deoxynivalenol in maize and 8 μg kg^{-1} of aflatoxin B$_1$ in peanuts corresponded to current EU regulatory limits. Final results, achieving a balance between classification accuracy and model stability (i.e., avoiding over-fitting), were obtained by thorough investigation of subsets based on available sample information. For maize samples, two different varieties (CRC and SBL) were available and separate models were created for both. While, for example, a CRC-only model resulted in a classification accuracy of 90%, the addition of SBL samples increased the sample

< 1750 µg/kg	17	17
> 1750µg/kg	6	70

Figure 3(a). Classification results for maize samples with contamination above and below the threshold of 1750 µg kg^{-1}; data labels provide deoxynivalenol concentration in µg kg^{-1}. The spectral window between 1800 and 800 cm^{-1} was used for chemometric modelling.

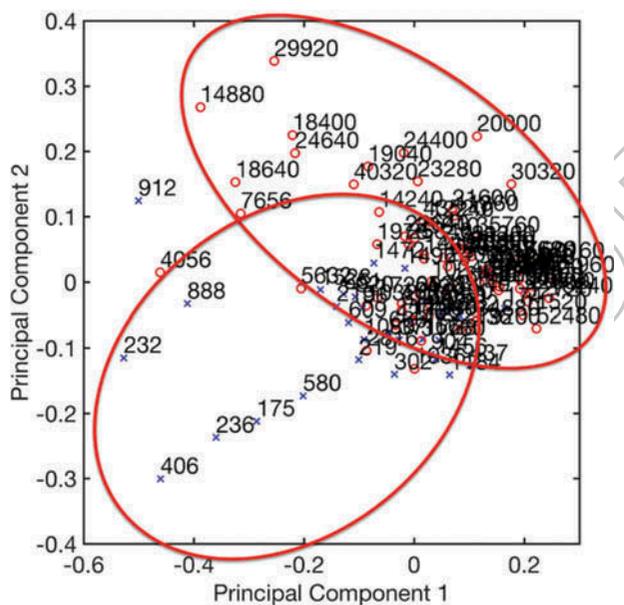|		
< 500 µg/kg	3	15
> 500 µg/kg	1	91

Figure 3(b). Classification results for maize samples with contamination above and below the threshold of 500 µg kg^{-1} using the decision tree method; data labels provide deoxynivalenol concentration in µg kg^{-1}. The spectral window between 1800 and 800 cm^{-1} was used for chemometric modelling.

Figure 3(c). Classification results for maize samples employing PCA score plots for principal components 1 and 2 using the 1750 µg kg^{-1} classification threshold with tentative assignment of clusters; data labels provide deoxynivalenol concentration in µg kg^{-1}. The spectral window between 1800 and 800 cm^{-1} was used for chemometric modelling.

number from 72 to 110, though there was degraded model performance (79%) compared with single-variety subsets, but not to unacceptable levels. The latter is preferable since more samples can be used for modelling, thus contributing to increased model stability. As a consequence, pooling of the two different available varieties was feasible, but we speculate that further

investigation and model revalidation will be required for addition of other varieties.

A similar approach investigated the impact of different fungal species present in the sample in model performance. While the presence of *F. graminearum* and *F. culmorum* did not have an impact on model performance, samples contaminated with *F. verticillioides* decreased classification accuracy only to 73%. Therefore, 74 samples infected with *F. verticillioides* were removed. A possible explanation could be that it has been reported that protein content alteration depends on the fungal species with which the commodity was infected (Boyacioğlu & Hettiarachchy 1995). Since protein bands were part of the spectral window used for chemometric modelling, it was feasible that the presence of *F. verticillioides* negatively impacted classification accuracy. For peanuts, which were sampled from African markets, additional sample information such as variety and fungal contamination was unavailable. The presence of aflatoxins within a large proportion of these peanut samples indicates widespread contamination by *Aspergillus* species.

Seventeen out of 110 samples in maize were classified as false-positives by the decision tree model. Note, too, that the model errs on the safe side. Also, only three contaminated samples were classified as below the defined threshold of 1750 µg kg^{-1}. This is in part due to the number of 23 low contamination samples compared with 87 contaminated ones. For peanuts, results of false-positives (10 out of 36 blanks; 28%) and false-negatives (11 out of 56 contaminated; 20%) were similar.

The concentration distribution of the target analyte in the model training set was important (Figures 4(a) and (b)). A large training set with a wide range of concentrations was essential for classification at the desired concentration level. Five out of 110 maize samples were close to the regulatory threshold with deoxynivalenol concentrations between 1500 and 2000 µg kg^{-1}, 32 samples contained less than 1500 µg kg^{-1}, and 73 samples more than 2000 µg kg^{-1}. For peanuts, aflatoxin B$_1$ concentrations were between 4 and 12 µg kg^{-1} for seven samples; 34 samples were < 4 µg kg^{-1}. It is notable that a classification threshold of 500 µg kg^{-1} showed an improved classification accuracy of 85% after cross-validation. It has been pointed out that, ideally, the training set must be representative of the classes, and at the same time include the full range of variability for the class (Schowengerdt 2006). For the 500 µg kg^{-1} threshold the number of samples around the regulatory threshold with deoxynivalenol concentrations between 250 and 750 µg kg^{-1} was larger (12 out of 110 maize compared with five for the 1750 µg kg^{-1} threshold). Therefore, we assume that the variability

| < 8μg/kg | 26 | 10 |
| > 8μg/kg | 11 | 45 |

Figure 4(a). Classification results for peanut samples with contamination above and below the threshold of 8 μg kg^{-1}; decision tree method; data labels provide aflatoxin B$_1$ concentration in μg kg^{-1}. The spectral window between 1800 and 800 cm^{-1} was used for chemometric modelling.
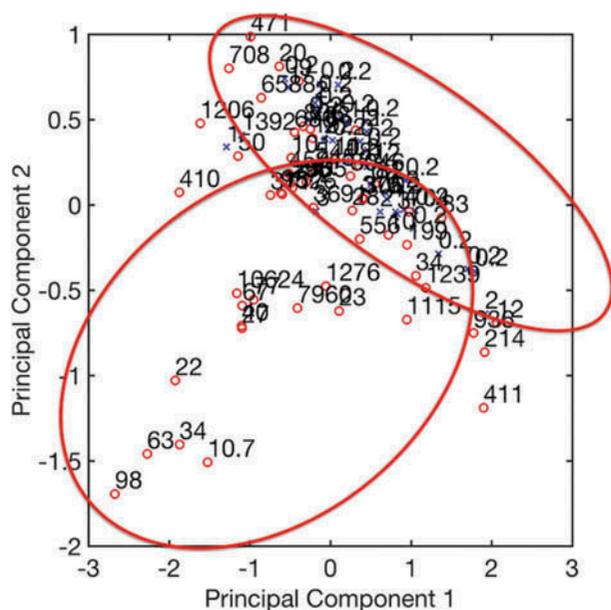
Figure 4(b). Classification results for peanut samples using PCA with tentative assignment of clusters; data labels provide aflatoxin B$_1$ concentration in μg kg^{-1}. The spectral window between 1800 and 800 cm^{-1} was used for chemometric modelling.

around the threshold was better defined, resulting in a small performance increase. As a result, samples close to the regulatory threshold are important to be part of the training set in order to model small differences at the defined classifier concentration effectively. These should also cover the natural variability unrelated to fungal contamination in a representative fashion. Subsets with large concentration differences between samples, i.e., a large number of both blanks and highly contaminated samples without samples at the regulatory threshold, yielded impressive classification results for the training set. Poor results for samples with toxin levels at the regulatory threshold were observed, though.

Previous work (Table 2) attempting classification of deoxynivalenol contaminated samples in maize only employed a small number of 15 samples for classification, although a sample with a concentrations as low as

309 μg kg^{-1} was classified as contaminated compared with true (i.e., below the reference method LOD) blanks (Kos et al. 2003). Regarding other matrices, the IR method presented by Abramovic et al. (2007) employed 17 samples (with three blanks < LOD) with a concentration range higher (i.e., 2.51–12.14 mg kg^{-1}), making it impossible to attain the 1750 and 500 μg kg^{-1} classification thresholds presented here. Deoxynivalenol in wheat of a single genotype was determined. The most recent report by De Girolamo et al. (2014) using an linear discriminant analysis (LDA) classification model for wheat NIR data showed similar performance with 70–90% of samples correctly classified. The classifier for the smallest concentration was set to < 1000 μg kg^{-1}. The sample set was large with 232 out of 464 samples used for model training. The authors noted an issue with the deoxynivalenol concentration distribution as a limiting factor for model stability for some of the multivariate methods tested. For peanuts this report is the first attempt to classify successfully aflatoxin B$_1$ at legal limits. Previous reports focused on quantification models making direct comparison difficult. Work by Fernandez-Ibanez et al. (2009) investigated 152 samples and defined samples with an aflatoxin B$_1$ concentration < 20 μg kg^{-1} as blanks. Qiang et al. (2014) investigated 82 rice samples using the same 20 μg kg^{-1} lower limit for contaminated samples. Finally, with the same lower limit Lee et al. (2015) tested 232 maize samples with concentrations up to 1206 μg kg^{-1}. This lower limit for contaminated samples were high and outside the EU legal limits compared with the 8 μg kg^{-1} used in the presented work.

The number of samples this study was an important step forward compared with previous work. Significant challenges remain to obtain well-defined, representative datasets with a suitable toxin concentration distribution. While this constitutes a considerable development effort, practical application is straightforward once the robustness of a model (e.g., for use with several varieties) has been established. Furthermore, 'self-learning' reference sets could be created by adding newly classified samples to the database after verification by LC-MS measurements. Thus, development of tailored models for well-defined sample sets, e.g., where variety and fungal contamination are known, becomes feasible.

Conclusions

The presented bagged decision tree method was able to classify maize samples at the 1750 and 500 μg kg^{-1} thresholds for deoxynivalenol with an accuracy of

79% and 85% respectively, and peanut samples for aflatoxin B$_1$ at 8 µg kg^{-1} with an accuracy of 77%, thus demonstrating the method's capabilities to classify at the respective regulatory limits. For aflatoxin B$_1$ in peanuts, classification at these low concentration levels was achieved for the first time. Minimal sample preparation, speedy data acquisition and automatic data analysis provide a measurement platform suitable for use in the field, e.g., for testing at points of entry by importers and during storage and processing. It could be used to provide a fast overview of an imported batch after preliminary characterisation (i.e., determination of variety and fungal species present), and to provide a decision-making tool for importers, if a more thorough investigation by LC-MS is required. The method can be implemented without the need for highly qualified personnel on-site. The recent occurrence of aflatoxins in European maize samples underlines the necessity and urgency for rapid, cheap mycotoxin detection methods available to food producers and importers as a changing climate potentially creates new challenges.

Acknowledgements

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

ORCID

Boris Mizaikoff http://orcid.org/0000-0002-5583-7962

References

Abramovic B, Jajic I, Abramovic B, Cosic J, Juric V. 2007. Detection of deoxynivalenol in wheat by Fourier transform infrared spectroscopy. Acta Chim Slov. 54:859.

Al-Jowder O, Kemsley EK, Wilson RH. 1997. Mid-infrared spectroscopy and authenticity problems in selected meats: a feasibility study. Food Chem. 59:195–201.

Amiel C, Mariey L, Curk-Daubie MC, Pichon P, Travert J. 2000. Potentiality of Fourier transform infrared spectroscopy (FTIR) for discrimination and identification of dairy lactic acid bacteria. Le Lait. 80:445–459.

Boyacioğlu D, Hettiarachchy NS. 1995. Changes in some biochemical components of wheat grain that was infected with *Fusarium graminearum*. J Cereal Sci. 21:57–62.

Breiman L. 1996. Bagging predictors. Mach Learn. 24:123–140.

Cerna M, Barros AS, Nunes A, Rocha SM, Delgadillo I, Copikova J, Coimbra MA. 2003. Use of FT-IR spectroscopy as a tool for the analysis of polysaccharide food additives. Carbohydr Polym. 51:383–389.

Chauhan R, Singh J, Sachdev T, Basu T, Malhotra BD. 2016. Recent advances in mycotoxins detection. Biosens Bioelectron. 81:532–545.

Choudhary AK, Kumari P. 2010. Management of mycotoxin contamination in pre-harvest and post-harvest crops: present status and future prospects. J Phytol. 2.

Chungu C, Mather D, Reid L, Hamilton R. 1996. Comparison of techniques for inoculating maize silk, kernel, and cob tissues with *Fusarium graminearum*. Plant Dis. 80:81–84.

Cozzolino D, Kwiatkowski M, Parker M, Cynkar W, Dambergs R, Gishen M, Herderich M. 2004. Prediction of phenolic compounds in red wine fermentations by visible and near infrared spectroscopy. Anal Chim Acta. 513:73–80.

Cozzolino D, Moron A. 2004. Exploring the use of near infrared reflectance spectroscopy (NIRS) to predict trace minerals in legumes. Anim Feed Sci Technol. 111:161–173.

De Girolamo A, Cervellieri S, Visconti A, Pascale M. 2014. Rapid analysis of deoxynivalenol in durum wheat by FT-NIR spectroscopy. Toxins. 6:3129–3143.

De Girolamo A, Lippolis V, Nordkvist E, Visconti A. 2009. Rapid and non-invasive analysis of deoxynivalenol in durum and common wheat by Fourier-transform near infrared (FT-NIR) spectroscopy. Food Addit Contam Part A. 26:907–917.

Draganova T, Daskalov P, Tsonev R. 2010. An approach for identifying of *Fusarium* infected maize grains by spectral analysis in the visible and near infrared region, SIMCA models, parametric and neural classifiers. Int J Bioautomation. 14:119–128.

Ertas N, Gonulalan Z, Yildirim Y, Karadal F. 2011. A survey of concentration of aflatoxin M1 in dairy products marketed in Turkey. Food Control. 22:1956–1959.

European Commission. 2002. Commission decision of 12 August 2002 implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results. Off J Eur Union. L221:9–36.

Fernandez-Ibanez V, Soldado A, Martinez-Fernandez A, De La Roza-Delgado B. 2009. Application of near infrared spectroscopy for rapid detection of aflatoxin b1 in maize and barley as analytical quality assessment. Food Chem. 113:629–634.

Fernández Pierna JA, Volery P, Besson R, Baeten V, Dardenne P. 2005. Classification of modified starches by Fourier transform infrared spectroscopy using support vector machines. J Agric Food Chem. 53:6581–6585.

Firrao G, Torelli E, Gobbi E, Raranciuc S, Bianchi G, Locci R. 2010. Prediction of milled maize fumonisin contamination by multispectral image analysis. J Cereal Sci. 52:327–330.

Groopman JD, Cain LG, Kensler TW, Harris CC. 1988. Aflatoxin exposure in human populations: measurements

and relationship to cancer. CRC Crit Rev Toxicol. 19:113–145.

IARC. 1993. Some naturally occurring substances: food items and constituents, heterocyclic aromatic amines and mycotoxins. Lyon: Working Group on the Evaluation of Carcinogenic Risks to Humans: International Agency for Research on Cancer.

IARC. 2016. Mycotoxin control in low- and middle-income countries.

Inon FA, Garrigues S, De La Guardia M. 2004. Nutritional parameters of commercially available milk samples by FTIR and chemometric techniques. Anal Chim Acta. 513:401–412.

Karoui R, Downey G, Blecker C. 2010. Mid-infrared spectroscopy coupled with chemometrics: a tool for the analysis of intact food systems and the exploration of their molecular structure–quality relationships – a review. Chem Rev. 110:6144–6168.

Kaya-Celiker H, Mallikarjunan PK, Kaaya A. 2016. Characterization of invasion of genus *Aspergillus* on peanut seeds using FTIR-PAS. Food Anal Methods. 9:105–113.

Kos G, Lohninger H, Krska R. 2003. Development of a method for the determination of *Fusarium* fungi on corn using mid-infrared spectroscopy with attenuated total reflection and chemometrics. Anal Chem. 75:1211–1217.

Lee K-M, Davis J, Herrman TJ, Murray SC, Deng Y. 2015. An empirical evaluation of three vibrational spectroscopic methods for detection of aflatoxins in maize. Food Chem. 173:629–639.

Lee KM, Herrman TJ. 2015. Determination and prediction of fumonisin contamination in maize by surface–enhanced Raman spectroscopy (SERS). Food Bioprocess Technol. 1.

Maertens K, Reyns P, De Baerdemaeker J. 2004. On-line measurement of grain quality with NIR technology. Trans ASAE. 47:1135–1140.

Malachová A, Beltran E, Sulyok M, Berthiller F, Krska R. 2015. Multi- toxin determination in food–the power of 'dilute and shoot' approaches in LC–MS–MS. LCGC Eur. 28.

Malachová A, Sulyok M, Beltran E, Berthiller F, Krska R. 2014. Optimization and validation of a quantitative liquid chromatography–tandem mass spectrometric method covering 295 bacterial and fungal metabolites including all regulated mycotoxins in four model food matrices. J Chromatogr A. 1362:145–156.

Mariey L, Signolle JP, Amiel C, Travert J. 2001. Discrimination, classification, identification of microorganisms using FTIR spectroscopy and chemometrics. Vib Spectrosc. 26:151–159.

Mark H, Workman J Jr. 2010. Chemometrics in spectroscopy. New York, NY: Academic Press.

Mauer LJ, Chernyshova AA, Hiatt A, Deering A, Davis R. 2009. Melamine detection in infant formula powder using near-and mid-infrared spectroscopy. J Agric Food Chem. 57:3974–3980.

McMullin D, Mizaikoff B, Krska R. 2015. Advancements in IR spectroscopic approaches for the determination of fungal derived contaminations in food crops. Anal Bioanal Chem. 407:653–660.

Mesterházy À. 1977. Reaction of winter wheat varieties to four *Fusarium* species. J Phytopathol. 90:104–112.

Miller JD. 2008. Mycotoxins in small grains and maize: old problems, new challenges. Food Addit Contam. 25:219–230.

Naes T, Isaksson T, Fearn T, Davies T. 2002. A user friendly guide to multivariate calibration and classification. Chichester, UK: NIR publications.

Pestka J. 2010. Toxicological mechanisms and potential health effects of deoxynivalenol and nivalenol. World Mycotoxin J. 3:323–347.

Prandini A, Tansini G, Sigolo S, Filippi L, Laporta M, Piva G. 2009. On the occurrence of aflatoxin m1 in milk and dairy products. Food Chem Toxicol. 47:984–991.

Qiang Z, Fuguo J, Chenghai L, Jingkun S, Xianzhe Z. 2014. Rapid detection of aflatoxin b1 in paddy rice as analytical quality assessment by near infrared spectroscopy. Int J Agric Biol Eng. 7:127.

Quinlan JR. 1986. Induction of decision trees. Mach Learn. 1:81–106.

Refaeilzadeh P, Tang L, Liu H. 2009. Cross-validation. In: Encyclopedia of database systems. Springer; p. 532–538.

Richard JL. 2007. Some major mycotoxins and their mycotoxicoses–an overview. Int J Food Microbiol. 119:3–10.

Rodrigues I, Handl J, Binder EM. 2011. Mycotoxin occurrence in commodities, feeds and feed ingredients sourced in the middle east and Africa. Food Addit Contam Part B. 4:168–179.

Schowengerdt RA. 2006. Remote sensing: models and methods for image processing. Academic Press.

Sirisomboon CD, Putthang R, Sirisomboon P. 2013. Application of near infrared spectroscopy to detect aflatoxigenic fungal contamination in rice. Food Control. 33:207–214.

Smela ME, Currier SS, Bailey EA, Essigmann JM. 2001. The chemistry and biology of aflatoxin b1: from mutational spectrometry to carcinogenesis. Carcinogenesis. 22:535–545.

Steyerberg EW, Harrell FE, Borsboom GJ, Eijkemans M, Vergouwe Y, Habbema JDF. 2001. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. J Clin Epidemiol. 54:774–781.

Swamy H, Smith TK, Karrow NA, Boermans HJ. 2004. Effects of feeding blends of grains naturally contaminated with *Fusarium* mycotoxins on growth and immunological parameters of broiler chickens. Poult Sci. 83:533–543.

Turner NW, Subrahmanyam S, Piletsky SA. 2009. Analytical methods for determination of mycotoxins: a review. Anal Chim Acta. 632:168–180.

Vardon P, McLaughlin C, Nardinelli C. 2003. Potential economic costs of mycotoxins in the United States. p. 139.

Vermeulen P, Pierna JAF, Egmond HPV, Dardenne P, Baeten V. 2012. Online detection and quantification of ergot bodies in cereals using near infrared hyperspectral imaging. Food Addit Contam Part A. 29:232–240.

Williams PC, Sobering DC. 1993. Comparison of commercial near infrared transmittance and reflectance instruments for analysis of whole grains and seeds. J Near Infrared Spectrosc. 1:25–32.

Williams PJ, Geladi P, Britz TJ, Manley M. 2012. Investigation of fungal development in maize kernels using NIR hyperspectral imaging and multivariate data analysis. J Cereal Sci. 55:272–278.

Wu F. 2007. Measuring the economic impacts of *Fusarium* toxins in animal feeds. Anim Feed Sci Technol. 137:363–374.

Young C, Kim S-S, Luzinova Y, Weida M, Arnone D, Takeuchi E, Day T, Mizaikoff B. 2009. External cavity widely tunable quantum cascade laser based hollow waveguide gas sensors for multianalyte detection. Sens Actuators B Chem. 140:24–28.