Food Chemistry 209 (2016) 286–292

Contents lists available at ScienceDirect

Food Chemistry

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

Quantitative determination of aflatoxin B1 concentration in acetonitrile by chemometric methods using terahertz spectroscopy

Hongyi Ge [a,b], Yuying Jiang [c,d], Feiyu Lian [a,b], Yuan Zhang [b,*], Shanhong Xia [c]

[a] College of Information Science and Engineering, Henan University of Technology, Zhengzhou, Henan 450001, China
[b] Key Laboratory of Grain Information Processing and Control, Ministry of Education, Zhengzhou 450001, China
[c] State Key Laboratory of Transducer Technology, Institute of Electronics, Chinese Academy of Sciences, Beijing 100080, China
[d] University of the Chinese Academy of Sciences, Beijing 100080, China

ARTICLE INFO

Article history:
Received 30 January 2015
Received in revised form 27 March 2016
Accepted 18 April 2016
Available online 20 April 2016

Keywords:
Aflatoxin B1
Quantitative analysis
Terahertz spectroscopy
PCR
PLS
SVM

ABSTRACT

Aflatoxins contaminate and colonize agricultural products, such as grain, and thereby potentially cause human liver carcinoma. Detection via conventional methods has proven to be time-consuming and complex. In this paper, the terahertz (THz) spectra of aflatoxin B1 in acetonitrile solutions with concentration ranges of 1–50 μg/ml and 1–50 μg/l are obtained and analyzed for the frequency range of 0.4–1.6 THz. Linear and nonlinear regression models are constructed to relate the absorption spectra and the concentrations of 160 samples using the partial least squares (PLS), principal component regression (PCR), support vector machine (SVM), and PCA-SVM methods. Our results indicate that PLS and PCR models are more accurate for the concentration range of 1–50 μg/ml, whereas SVM and PCA-SVM are more accurate for the concentration range of 1–50 μg/l. Furthermore, ten unknown concentration samples extracted from mildewed maize are analyzed quantitatively using these methods.

1. Introduction

Aflatoxins are mycotoxins produced by the fungus *Aspergillus flavus* and the mold *Aspergillus parasiticus*. Aflatoxins are toxic to humans and are among the most carcinogenic substances known (Abrar et al., 2013). Aflatoxins are widespread in agricultural and food products such as groundnuts, maize, wheat, rice, and other dried foods; thus, they pose a potential threat to the safe consumption of agricultural products. Aflatoxins are associated with both acute and chronic toxicity in humans. Since they significantly contribute to the loss of agricultural products (particularly after harvests) and are detrimental to human health (Reddy et al., 2010), aflatoxins have received considerable attention as a topic of research. Aflatoxins are classified into several types, and the types of interest are aflatoxins B1, B2, M1, and M2. In addition, aflatoxin B1 (AFB1) is quite commonly found in agricultural products, and it is considered the most toxic. Thus, there is a need for rapid and reliable detection methods to achieve quantitative analysis of AFB1.

Several techniques have been employed to detect and quantify AFB1, such as thin-layer chromatography (TLC), high-performance liquid chromatography (HPLC), and enzyme-linked immunosorbent assays (ELISAs) (Turner, Subrahmanyam, & Piletsky, 2009). Although these methods are highly accurate, they are time-consuming (involving laborious preparation of samples) and complex, and hence, they cannot meet the demands of rapid detection. In recent years, spectroscopic techniques including near-infrared (NIR) spectroscopy (Fernandez-Ibanez, Soldado, Martinez-Fernandez, & de la Roza-Delgado, 2009) and Raman spectroscopy (Lee, Herrman, & Yun, 2014) have been used for the rapid detection and quantitative analysis of various compounds. However, these approaches cannot probe the far-infrared spectral region, which contains a wealth of physical and chemical information concerning the materials being investigated.

THz radiation (radiation in the frequency range from 0.1 to 3 THz) lies in the far-infrared range. THz time-domain spectroscopy (THz-TDS) has recently been used as a new characterization technique to reveal interesting material properties. It has been successfully employed in diverse fields such as materials science (Ferguson & Zhang, 2002), biology (Jepsen, Moller, & Merbold, 2007), chemistry (Siegel, 2004), manufacturing quality control and detection (Gowen, O'Sullivan, & O'Donnell, 2012), and security (Melinger, Laman, & Grischkowsky, 2008). Several studies have focused on quantitative analysis using THz-TDS. Ueno, Rungsawang, Tomita, and Ajito (2006) quantitatively analyzed the THz absorption spectra of amino acids via linear regression methods. Nishikiori et al. (2008) used the partial least squares (PLS) method to analyze L- and DL-tartaric acid mixtures in the

* Corresponding author.
E-mail addresses: zhangyuan@haut.edu.cn, gucasrr@gmail.com (Y. Zhang).

http://dx.doi.org/10.1016/j.foodchem.2016.04.070
0308-8146/

THz range from 0.1 to 3 THz. Hua, Zhang, and Zhou (2010) used THz-TDS and regression methods to examine cyfluthrin *n*-hexane solutions in the concentration range of 1–20 μg/ml. Ma, Wang, and Li (2013) quantitatively analyzed a mixture of thiabendazole and polyethylene using various algorithms including PLS, interval PLS, back interval PLS, and moving window PLS (mwPLS), and the most accurate results were achieved by using mwPLS. The results obtained in these studies support THz-TDS as a powerful tool for the quantitative analysis of compounds. However, few studies have considered nonlinear regression methods and optimal model parameters for the regression models used in THz-TDS studies.

In this work, we use THz-TDS together with chemometric methods to obtain the absorption spectra of AFB1 in acetonitrile solution in the concentration ranges of 1–50 μg/ml and 1–50 μg/l to evaluate the possibility of quantitatively analyzing AFB1. We analyze the optimal parameters of the models used for measurement accuracy, and we compare the results of the linear and nonlinear regression models.

2. Experimental methods

2.1. Experimental setup

THz transmission spectra of AFB1 were obtained using a THz-TDS system based on photoconductive switches (Fig. S1a in Supporting information). A mode-locked Ti:sapphire laser, which provided 100 fs pulses at a wavelength of 800 nm with a repetition rate of 80 MHz, was used as the radiation source. The main beam was divided into two beams. One of the beams (the pump beam) was used to generate THz pulses, and the other (the probe beam) was used to detect THz pulses. The details of the THz-TDS system are explained in Han, Huang, and Zhang (2000) and Stoik, Bohn, and Blackshire (2008). In order to avoid absorption of moisture by ambient air, the terahertz path was purged continuously with dry nitrogen gas, and the humidity was maintained at a constant value of 1% during the experiments. The effectively frequency range for this experiment system is 0.2–1.6 THz, and its peak dynamic range is >1000, with a signal-to-noise ratio of >5000.

2.2. Sample preparation

The standard solution of AFB1 in acetonitrile solution was purchased from Pribolab Co. Ltd. (Singapore) at a concentration of 100 ± 0.1 μg/ml; 20 types of solution were diluted to concentration ranges of 1–50 μg/ml and 1–50 μg/l from a stabilized solution respectively, for each type of solution, 4 samples were used without further processing, and these diluted 160 solution samples were kept at 4 °C until use. The exact solution amount of each sample was measured using a spectrophotometer (Samuel, Sivaramakrishna, & Mehta, 2014). Each sample was measured at a fixed layer by using a liquid cell with a Teflon spacer and polyethylene windows, the thickness of the solution layer was 100 μm, and the cell window was 1.5 mm. The liquid cell has higher transmission and lower absorption. The schematic of the liquid cell is shown in Fig. S1b (Supporting information). To ensure that no macroscopic bubbles existed in the solution, as their presence can negatively affect the results, the sample was inspected visually. Depending on the AFB1 concentration, the absorption index of each sample was different. All subsequent measurements were performed at room temperature.

2.3. Data acquisition

For each measurement, the stability of the system was evaluated with respect to a reference (containing no AFB1) and the sample under investigation. And each measured sample was compared with a reference sample; the waveforms of samples are similar with others, showing the stability of the system performance. The time-domain spectra of the sample and reference spectra were recorded by measuring the sample and reference, respectively. Subsequently, the frequency-domain power $E(\omega)$ of the THz pulse was obtained via the fast Fourier transform of the time-domain spectra (Arora et al., 2012). The frequency-domain spectra of the sample, the reference spectra, and the nitrogen are denoted as $E_S(\omega)$, $E_{ref}(\omega)$, and $E_n(\omega)$, respectively. Further, the Fabry–Perot effect of the liquid cell windows has not been taken into account; in the equations below, we use the numbers 1, 2, and 3 to represent the media of nitrogen, the empty liquid cell, and the sample, respectively. By using the equations that govern the propagation of light, $E_S(\omega)$, $E_{ref}(\omega)$, and $E_n(\omega)$, can be obtained (Duvillaret, Garet, & Coutaz, 1996) in the Supporting information.

The details of data acquisition were described in the Supporting Information. According to Eqs. (2) and (3) in Supporting information, the absorption coefficient $\alpha(\omega)$ and refractive index n of the sample can be calculated using the following equation (Liu, Yue, Wang, Sun, & Zhang, 2012):

$$n_2(\omega) = \frac{\arg[H_{Measure}(\omega)]c}{\omega d_2} + n_0 \tag{1}$$

$$\alpha_2 = \frac{2}{d_2} \ln\left[\frac{n_2(n_1+n_0)^2}{|H_{Measure}(\omega)|(\tilde{n}_1+n_2)^2 n_0}\right] \tag{2}$$

2.4. Modeling methods

To quantitatively analyze AFB1 concentrations in acetonitrile, we employed linear and nonlinear methods including PLS, principal component regression (PCR), support vector machine (SVM), and PCA-SVM. Further, the root mean square error (RMSE) and correlation coefficient (R) were applied to evaluate the performance of the proposed model (Hua et al., 2010; Ma et al., 2013). Depending on the values of R (higher) and RMSE (lower), we can obtain an improved model performance, which leads to improved prediction accuracy for the quantitative measurement of the unknown solution sample.

PLS and PCR are the most commonly used linear regression methods. The aim of regression procedure is to reduce the dimensionality of a set of possibly correlated variables while retaining maximum variability based on the variance-covariance structure, which is a linear combination of the original dataset (Katrin et al., 2012). Many more detailed descriptions of PCR have been published; some such descriptions can be found in Burnett et al. (2009) and Noori, Karbassi, and Sabahi (2010).

PLS (Brereton, 2000) is used to determine the linear decomposition between two matrices, input matrix and output matrix. It extracts the orthogonal characteristics from the sample spectra, then constructs the relationship between the characteristics and the concentration of samples. More details about PLS can be found in Blanco, Coellho, Iturriaga, Masposch, and Pages (2000) and Ham, Kostanic, Cohen, and Gooch (1997).

The SVM is a supervised learning method that analyzes data and recognizes patterns, and it is used for data classification and for making various types of predictions. A detailed description of the theory of SVM has been reported in He, Yang, and Xie (2013) and Cortes and Vapnik (1995). The details of these three multivariate analysis techniques in our present work were described in the Supporting Information.

In our study, we used the radial basis function (RBF) given by Eq. (3) for the SVM model for the following reasons: first, the linear kernel is a special case of the RBF; second, the RBF kernel has fewer

parameters than polynomial and sigmoid kernels do. Furthermore, RBF can produce good performance under small numbers of samples (Noori, Abdoli, Ghasrodashti, & Ghazizade, 2009)

$$k(x_i, y_i) =< \phi(x_i), \phi(x_j) >= \exp\left(-\frac{\|x_i - y_i\|^2}{\gamma^2}\right) \quad (3)$$

Here, x_i and y_i denote the input vectors, and γ denotes the RBF parameter.

The performance of different models constructed is compared in terms of the correlation coefficient between the reference and predicted value (R) and the root-mean-square error (RMSE) of the calibration set (RMSEC)

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(y_r^i - y_p^i)^2}{n}} \quad (4)$$

$$R = \frac{\sum_{i=1}^{n}(y_r^i - \overline{y_r})(y_p^i - \overline{y_p})}{\sqrt{\sum_{i=1}^{n}(y_r^i - \overline{y_r})^2 \sum_{i=1}^{n}(y_p^i - \overline{y_p})^2}} \quad (5)$$

where n represents the number of samples in the training set; y_r^i and y_p^i are the reference value of the ith sample in the data set and the predicted value of the ith sample in the developed model, respectively; and $\overline{y_i}$ and $\overline{y_p}$ are the average values of the reference values and the predicted values of the samples, respectively.

A set of 40 types of solution (160 samples in total) was used in this experiment. All the samples were divided into two sets randomly, the calibration set (80 samples) and the validation set (80 samples). The whole modeling processes of linear and nonlinear are shown in Fig. S2a (Supporting Information). The absorbance values of an appropriate frequency range are used as multidimensional input. PCA transforms input data into a set of values ofl inearly uncorrelated variables with a lower dimension. This can represent the original data and then uses extracted features from absorption spectra as the input of PCA-SVM model. The linear and nonlinear models can be constructed based on the input values from the calibration procedure. Furthermore, the aim of the prediction set of samples is to predict the concentration values using the built models, the model prediction process is shown in Fig. S2b (Supporting Information). Similarly, the absorption spectra of sample of prediction set are used for the four models. The predicted concentration of the sample can be obtained by the built models and the absorbance values.

3. Results and discussion

3.1. Experimental results

The absorption coefficients of AFB1 in acetonitrile solution were measured via THz-TDS in the frequency range 0.4–1.6 THz. Each concentration sample and reference was scanned four times, and the spectral records were averaged to generate a spectrum. The THz pulses corresponding to the sample and reference were obtained by scanning the liquid cell filled with AFB1 solution and pure acetonitrile solvent, respectively. In order to minimize complex baseline fluctuations, all spectra in the frequency domain were further processed with a standard normal variate algorithm (Barnes, Dhanoa, & Lister, 1989). Comparisons of the absorption coefficients of the different samples are shown in Fig. 1(a) (for the concentration range 1–50 μg/ml) and Fig. 1(b) (for the concentration range 1–50 μg/l). In order to represent the variation in measurement effectually, the average absorption coefficients of different concentrations range are shown in Fig. S3 (Supporting Information). And the average absorption coefficients of sample are 34.5114 (2 μg/l), 33.81602(4 μg/l), 30.79999(10 μg/l), 28.37413(15 μg/l), 25.94679(20 μg/l), 23.96259(25 μg/l), 23.67114(30 μg/l) and 20.72576(40 μg/l) respectively (for the concentration range 1–50 μg/l), and 80.88223 (1 μg/ml), 79.53026 (3 μg/ml), 78.16886(5 μg/ml), 75.91949(10 μg/ml), 74.29529 (15 μg/ml), 72.551(20 μg/ml), 71.34436(25 μg/ml), 67.44381 (30 μg/ml) and 65.58572(40 μg/ml) respectively (for the concentration range 1–50 μg/ml).

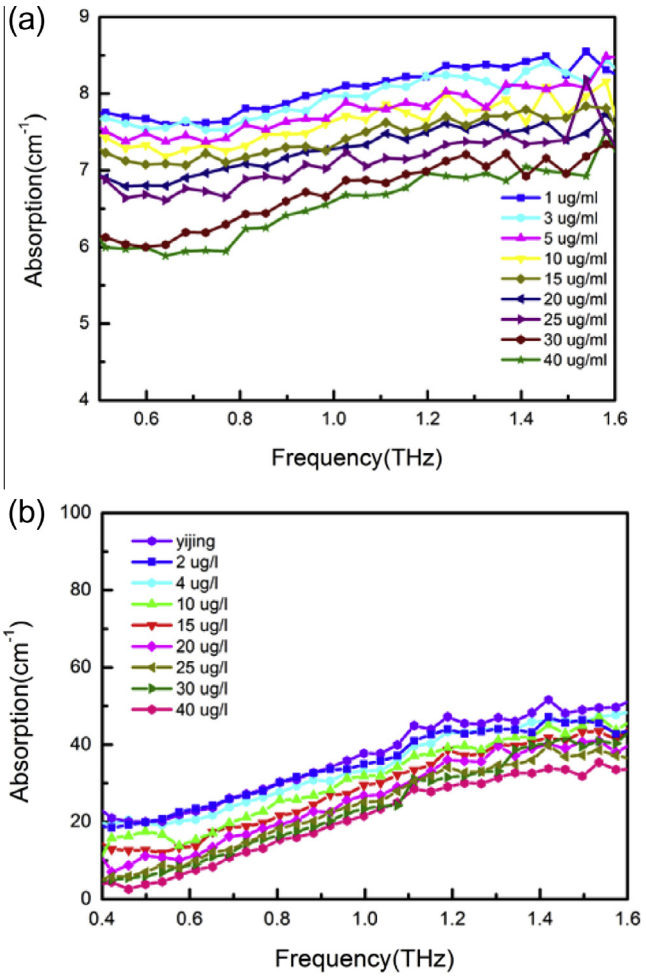

Fig. 1. Absorption coefficients of aflatoxin B1 (AFB1) at different concentrations as determined by terahertz time-domain spectroscopy (THz-TDS) for AFB1 concentration ranges of (a) 1–50 μg/ml and (b) 1–50 μg/l.

We observed a decrease in the absorption coefficient with increasing AFB1 concentrations over different concentration ranges. Since the AFB1 concentration increases, strong absorption of acetonitrile molecules is replaced by strong absorption of AFB1 molecules. This phenomenon indicates that THz absorption of acetonitrile solvent is considerably higher than that of solute AFB1. Here, we recall that THz radiation is sensitive to polar materials such as water (Arora et al., 2012). From Fig. 1, it is obvious that a featureless THz absorption increases with frequency. As regards sample concentrations ranging from 1 to 50 μg/ml (Fig. 1(a)), the waveform changes can be more easily distinguished than in the case of sample concentrations ranging from 1 to 50 μg/l (Fig. 1 (b)). The overall trend of the absorption coefficient curve in Fig. 1 (b) is a nearly linear increase with increasing frequency, whereas in Fig. 1(a), the curve is gentler. The spectra of AFB1 samples at higher frequencies (above 1.5 THz) produce lower SNR due to the limitation of the dynamic range of the measurement system, and the absorption spectra for samples with different concentrations can be seen to intersect, which could be introduced by the

Fabry-Perot effect of the cell walls. For the samples with lower concentrations, due to the irrotational bonding of a part of the molecule, the vibration and rotation of the molecule may be suppressed (Asaki, Redondo, Zawodzinski, & Taylor, 2002). Thus, a more intersection of the THz absorption spectra is to be expected in the 1–50 μg/l range.

In order to further interpret the variation of the absorption spectra of the solution samples, THz absorption spectra of AFB2 in acetonitrile solution are also measured by THz-TDS. Fig. 2 shows the absorption coefficients of AFB1 and AFB2 solution samples with concentrations of 10 μg/l and 30 μg/l. In order to represent the variation in measurement effectually, the average absorption coefficients of AFB1 and AFB2 solution samples with concentrations of 10 μg/l and 30 μg/l are shown in Fig. S4 (Supporting information). the average absorption coefficients of samples are 307.9999 (AFB1_10 μg/l), 328.6889(AFB2_10 μg/l), 236.7114(AFB1_30 μg/l), and 279.5138(AFB2_30 μg/l) respectively. It is clear that the same trend is obtained; however, the spectra between the two samples have obvious differences that indicate that the effect arises from a minor difference in molecular structure.

The waveforms indicate the response of aflatoxin molecules in the THz region. Furthermore, no characteristic absorption peaks are observed in the THz frequency range, and therefore, chemometric methods are employed to quantitatively analyze the AFB1 solution. Attempts were made to correlate the measured THz spectra of AFB1 with the corresponding concentrations by using regression methods.

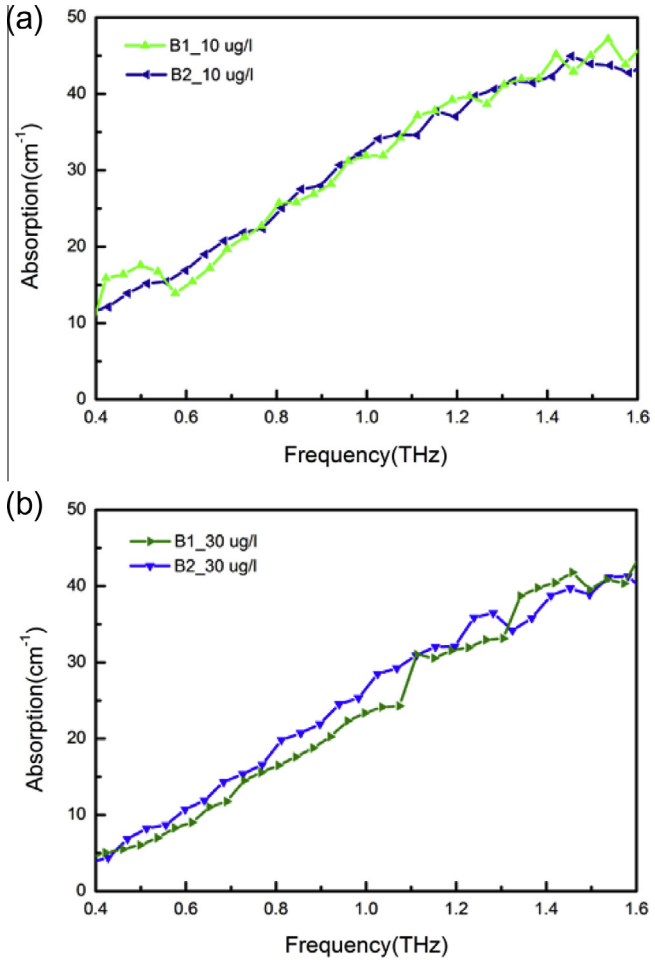

Fig. 2. Absorption coefficient of aflatoxin B1 (AFB1) and aflatoxin B2 (AFB2) at different concentrations as determined by terahertz time-domain spectroscopy (THz-TDS): (a) 10 μg/l and (b) 30 μg/l.

The quantitative analysis methods (PLS, PCR, SVM, and PCA-SVM) introduced in the previous section were applied to construct models relating the THz absorption coefficient and the concentrations of 160 samples. The multivariate calibrations were performed on THz spectral data in MATLAB software package (Version 2012a, Mathworks Inc., Natic, USA) using toolbox and user written scripts. A set of 160 samples was used in this experiment. All the samples were divided into two sets randomly with a ratio of 1:1, the calibration set (80 samples) and the validation set (80 samples). The absorption spectra corresponding to the two above-mentioned concentration ranges were used as the input to the models. An appropriate frequency range (0.8–1.4 THz) is used, the 4 factors for PLS, 5 factors for PCR and 4 factors for PCA-SVM are used according to the least RMSE, results of the calibration and the validation sets for sample concentrations ranging from 1 to 50 μg/ml are listed in Table 1.

It can be seen from Table 1 that the results of the linear regression models (PLS and PCR) are more accurate than those of the nonlinear models (SVM and PCA-SVM) based on the lower RMSE and higher R values. There is no obvious difference between the performances of PLS and PCR, whereas the SVM and PCA-SVM do not provide satisfactory results. Scatter plots of the predicted concentration versus the reference concentration with linear and nonlinear methods are shown in Fig. 3.

As can be seen in Fig. 3(a), the distributions of points for the concentrations predicted by the PLS and PCR models lie closer to the reference line, as compared to those predicted by the SVM and PCA-SVM models. However, in Fig. 3(b), the concentrations predicted by the SVM and PCA-SVM models are higher than those of the PLS and PCR models.

In order to further compare the performances of the prediction models, we analyzed the predicted concentrations of the samples using the models mentioned. The performances of the models are indexed by the prediction accuracies, as shown in Table 2.

From Table 2, we observe that the prediction results of the models as a whole show that the performances of the models were satisfactory for the determination of AFB1. Comparing the results obtained for each model, we note that the prediction accuracies of different models differ greatly between the two concentration ranges. The results show that the accuracies of the SVM and PCA-SVM predictions were poor in the concentration range of 1–50 μg/ml, whereas the prediction accuracies of the PLS and PCR models were as high as 87.5%. For the concentration range of 1–50 μg/l, the prediction accuracies of SVM and PCA-SVM reach 93.75%, whereas the PLS and PCR performances were relatively poor.

We note that the general trend related to prediction accuracies is different for the two cases, and this can be expressed as follows: PCR > PLS > SVM > PCA-SVM (1–50 μg/ml) and PCA-SVM > SVM > PCR > PLS (1–50 μg/l).

In order to further demonstrate the feasibility of the proposed method for detecting AFB1 in agricultural products, AFB1 in maize was analyzed to obtain predicted concentration values. Generally, AFB1 exists in solid form in maize; thus, it needs to be extracted

Table 1
Comparison of different models for calibration and validation data sets.

Model	Frequency	Factors	Calibration		Validation	
			R	RMSE (μg/ml)	R	RMSE (μg/ml)
PLS	0.8–1.4 THz	4	0.985	0.753	0.983	0.691
PCR	0.8–1.4 THz	5	0.986	0.587	0.982	0.643
SVM	0.8–1.4 THz		0.967	1.365	0.963	1.674
PCA-SVM	0.8–1.4 THz	4	0.953	1.864	0.947	1.953

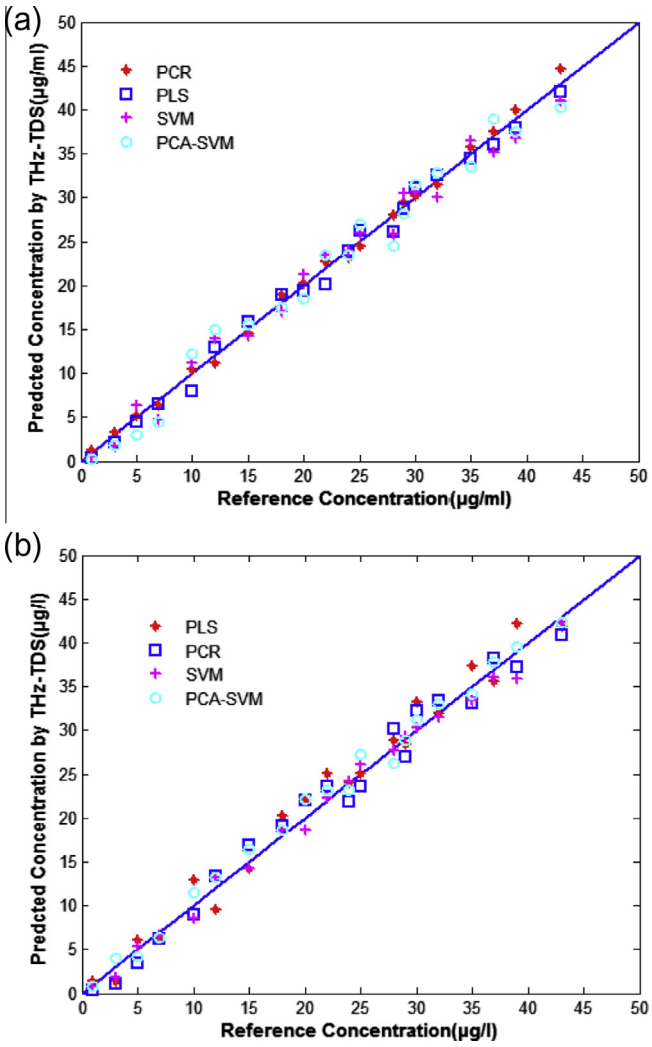

Fig. 3. Plots of the reference concentration versus the aflatoxin B1 (AFB1) concentration predicted by terahertz time-domain spectroscopy (THz-TDS) based on partial least squares (PLS), principle component analysis (PCR), support vector machine (SVM), and PCA-SVM models calculated in the spectral range from 0.8 to 1.4 THz for (a) concentrations ranging from 1 to 50 μg/ml and (b) concentrations ranging from 1 to 50 μg/l.

Table 2
Results of prediction accuracies using different analysis methods.

Model	Prediction accuracies (%)	
	1–50 μg/ml	1–50 μg/l
PLS	82.5%	35%
PCR	87.5%	50%
SVM	45%	85%
PCA-SVM	37.5%	93.75%

by traditional methods, such as by using immunoaffinity columns (the details of the extraction process can be found in Xie et al. (2014)) and preserved in a polar solvent (acetonitrile). In the test, ten samples of AFB1 solution are prepared for quantitative analysis using PLS, PCR, SVM, and PCA-SVM methods. Each sample was measured four times to generate an average spectrum, where the scan time for each sample was no more than 1 s. Thus, the prediction time for each sample is no more than 5 s (including the total scan time of 4 s and the calculation time of 1 s) on a standard PC with 2 GB of RAM and a Pentium CPU. The prediction results of different models are shown in Table 3.

From Table 3, it is clear that the PCA-SVM and SVM methods can 1) obtain the best prediction accuracy in the concentration range of 1–50 μg/l and 2) attain the minimum prediction error, indicating that the prediction results of the unknown concentration samples extracted from the maize are in agreement with the results based on standard AFB1 samples using the four regression model methods. Moreover, we observe that the prediction results of sample numbers 1, 5, and 10 calculated by PLS are both 0 μg/l, and the prediction result of sample 1 and 2 are 0 μg/l by PCR, showing that the PLS and PCR methods cannot attain the appropriate prediction value in the lower concentration range due to the lower sensitivity of the system, the noise, and the decrease of the linear relation between the absorption spectra and the concentrations of the samples.

In addition, the minimum prediction errors of the ten samples are all less than 3 μg/l, indicating that the prediction results of the PCA-SVM and SVM models can obtain more accurate measurements. The prediction errors may be attributed the cases that the reference values of testing set are affected by the operator. Further work for prediction models is greatly recommended as the results of models are not enough to satisfy the requirement of optimal values. This result also demonstrates that the THz-TDS combined with chemometric methods is a novel potential detection tool for quantitative analysis of AFB1 in food and agricultural products.

3.2. Discussion

Traditional methods for detecting AFB1 are mostly time-consuming and complex. Whereas TLC and HPLC (Shim et al., 2007) have been officially adopted for the determination of AFB1, ELISA is easier to use compared with chromatographic methods and has been applied increasingly for the development of analytical methods for AFB1; furthermore, TLC, HPLC, and ELISA typically require skilled operators, washing steps, and multiple incubations (Stroka & Anklam, 2002). Thus, these technologies are unsuitable for fast detection.

THz spectroscopy is a rapid and noninvasive detection method for detecting AFB1. In this paper, the THz absorption spectra of AFB1 were measured and analyzed. The PCR, PLS, SVM, and PCA-SVM regression methods were employed to construct models for the determination of AFB1. The unknown concentration samples extracted from mildewed maize were applied to validate the feasibility of the models. The prediction process for each sample of unknown concentration can be completed within 5 s. The results indicate that the THz-TDS combined with chemometric methods can be used for the fast determination of AFB1, and may enable the possibility of determination of AFB1 in food and agricultural products.

Although the PCA improves the SVM operator, there have great potential for improvement the performance of models. Further research should be implemented for the improvement of the

Table 3
Prediction results using different models.

Sample number	PLS	PCR	SVM	PCA-SVM	Ref. (μg/l)	Min. Error
1	0	0	9.59	8.55	7.15	1.40
2	7.43	0	9.35	12.91	11.27	1.64
3	10.99	9.85	15.79	15.32	14.23	1.09
4	12.57	11.62	13.79	15.20	16.23	1.03
5	0	18.54	30.91	29.24	27.13	2.11
6	34.69	35.97	32.14	32.35	29.75	2.39
7	38.33	31.40	32.8	33.12	34.93	1.81
8	48.17	27.92	40.36	34.97	36.7	1.73
9	45.29	50.63	40.91	42.7	41.67	0.76
10	0	40.59	43.29	43.76	45.3	1.54

model accuracy and the reduction of the threshold for the detection of samples. It will be necessary to take into account the experimental uncertainties, which include the system and background noise as well as the system sampling accuracy (Withayachumnankul, Fischer, Lin, & Abbott, 2008). Meanwhile, the sample preparation process is essential to the determination of AFB1. The error in the process of sample preparation is affected by the standard solution accuracy, the precision of the diluting process, and the preservation environment. To decrease the detection threshold, we suggest the investigation and development of 1) a more suitable pretreatment method for the absorption spectra (the frequency range selection of which can reflect the most relevant information together with the sample spectra; exploring the possibilities of each of the preprocessing methods and some of their combinations), 2) more accurate models, such as an appropriate frequency range as input of the constructed models, with more relevant information, which can cause less prediction error, and 3) suitable algorithms, in order to improve the measurement accuracy, multisource information fusion technique based on DS evidence theory method will be employed to the analyze quantitatively in the future research. Furthermore, complex situations, such as the existence of more than two components in the solution, also need to be considered to explore more effective methods for the determination of AFB1 in real samples.

4. Conclusion

We have demonstrated the feasibility of the quantitative determination of AFB1 in acetonitrile solution by chemometric methods using THz-TDS. The absorption spectra of AFB1 solution samples were obtained and analyzed in the frequency range of 0.4–1.6 THz. To quantitatively analyze the concentrations of the samples, the PLS, PCR, SVM, and PCA-SVM models were employed. Our experimental results indicate the superior performance of the linear regression models over the nonlinear regression models for the concentration range of 1–50 μg/ml, while for the concentration range of 1–50 μg/l, the nonlinear-regression-based models provide significantly more accurate results than those provided by the linear-regression-based models. In addition, samples of unknown concentrations are employed to perform quantitative analyses using the regression methods. This confirms that THz spectroscopy is a promising technique that provides a rapid and reliable method for quantitative toxin detection in agricultural products.

Acknowledgments

This work was supported by National High-tech R&D Program of China (863 Program) (Grant No. 2012AA101608) and the National Natural Science Foundation of China (Grant No. 61071197). Key Scientific and Research Project of Educational Committee of Henan Province of China (Grant No. 16A510002), Key Science and Technology Program of Henan Province of China (Grant No. 122102210217). The authors declare no conflict of interest.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.foodchem.2016.04.070.

References

Abrar, M., Anjum, F. M., Butt, M. S., Pasha, I., Randhawa, M. A., Saeed, F., & Waqas, K. (2013). Aflatoxins: Biosynthesis, occurrence, toxicity, and remedies. *Critical Reviews in Food Science and Nutrition, 53*, 862–874.

Arora, A., Luong, T. Q., Kruger, M., Kim, Y. J., Nam, C. H., Manz, A., & Havenith, M. (2012). Terahertz-time domain spectroscopy for the detection of PCR amplified DNA in aqueous solution. *Analyst, 137*, 575–579.

Asaki, M. L. T., Redondo, A., Zawodzinski, T. A., & Taylor, A. J. (2002). Dielectric relaxation and underlying dynamics of acetonitrile and 1-ethyl-3-methylimidazolium triflate mixtures using THz transmission spectroscopy. *Journal of Chemical Physics, 116*, 10377–10385.

Barnes, R. J., Dhanoa, M. S., & Lister, S. J. (1989). Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Applied Spectroscopy, 43*, 772–777.

Blanco, M., Coellho, J., Iturriaga, H., Masposch, S., & Pages, J. (2000). NIR calibration in non-linear systems: Different PLS approaches and artificial neural networks. *Chemometrics and Intelligent Laboratory Systems, 50*, 75–82.

Brereton, R. G. (2000). Introduction to multivariate calibration in analytical chemistry. *Analyst, 125*, 2125–2154.

Burnett, A. D., Fan, W. H., Upadhya, P. C., Cunningham, J. E., Hargreaves, M. D., Munshi, T., ... Davies, A. G. (2009). Broadband terahertz time-domain spectroscopy of drugs-of-abuse and the use of principal component analysis. *Analyst, 134*, 1658–1668.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning, 20*, 273–297.

Duvillaret, L., Garet, F., & Coutaz, J. L. (1996). A reliable method for extraction of material parameters in terahertz time-domain spectroscopy. *IEEE Journal of Selected Topics in Quantum Electronics, 2*, 739–746.

Ferguson, B., & Zhang, X. C. (2002). Materials for terahertz science and technology. *Nature Materials, 1*, 26–33.

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 Chemistry, 113*, 629–634.

Gowen, A. A., O'Sullivan, C., & O'Donnell, C. P. (2012). Terahertz time domain spectroscopy and imaging: Emerging techniques for food process monitoring and quality control. *Trends in Food Science & Technology, 25*, 40–46.

Ham, F. M., Kostanic, I. N., Cohen, G. M., & Gooch, B. R. (1997). Determination of glucose concentrations in an aqueous matrix from NIR spectra using optimal time-domain filtering and partial least-squares regression. *IEEE Transactions on Biomedical Engineering, 44*, 475–485.

Han, P. Y., Huang, X. G., & Zhang, X. C. (2000). Direct characterization of terahertz radiation from the dynamics of the semiconductor surface field. *Applied Physics Letters, 77*, 2864–2866.

He, M., Yang, G. L., & Xie, H. Y. (2013). A hybrid method to recognize 3D object. *Optics Express, 21*, 6346–6352.

Hua, Y. F., Zhang, H. J., & Zhou, H. L. (2010). Quantitative determination of cyfluthrin in n-hexane by terahertz time-domain spectroscopy with chemometrics methods. *IEEE Transactions on Instrumentation and Measurement, 59*, 1414–1423.

Jepsen, P. U., Moller, U., & Merbold, H. (2007). Investigation of aqueous alcohol and sugar solutions with reflection terahertz time-domain spectroscopy. *Optics Express, 15*, 14717–14737.

Katrin, S., Philippe, C. C., Reinald, B., Bert, M., Cora, H., & Jacqueline, R. (2012). Automatic selection of a representative trial from multiple measurements using Principle Component Analysis. *Journal of Biomechanics, 45*, 2306–2309.

Lee, K. M., Herrman, T. J., & Yun, U. (2014). Application of Raman spectroscopy for qualitative and quantitative analysis of aflatoxins in ground maize samples. *Journal of Cereal Science, 59*, 70–78.

Liu, C., Yue, L. Y., Wang, X. K., Sun, W. F., & Zhang, Y. (2012). Measurement of optical parameters of organic solvents by THz time-domain reflection spectroscopy. *Spectroscopy and Spectral Analysis, 32*, 1471–1475.

Ma, Y. H., Wang, Q., & Li, L. Y. (2013). PLS model investigation of thiabendazole based on THz spectrum. *Journal of Quantitative Spectroscopy and Radiative Transfer, 117*, 7–14.

Melinger, J. S., Laman, N., & Grischkowsky, D. (2008). The underlying terahertz vibrational spectrum of explosives solids. *Applied Physics Letters, 93*, 011102.

Nishikiori, R., Yamaguchi, M., Takano, K., Enatsu, T., Tani, M., de Silva, U. C., ... Kawase, M. (2008). Application of partial least square on quantitative analysis of L-, D-, and DL-tartaric acid by terahertz absorption spectra. *Chemical & Pharmaceutical Bulletin, 56*, 305–307.

Noori, R., Abdoli, M. A., Ghasrodashti, A. A., & Ghazizade, M. J. (2009). Prediction of municipal solid waste generation with combination of support vector machine and principal component analysis: a case study of Mashhad. *Environmental Progress and Sustainable Energy, 28*, 249–258.

Noori, R., Karbassi, A. R., & Sabahi, M. S. (2010). Evaluation of PCA and Gamma test techniques on ANN operation for weekly solid waste prediction. *Journal of Environmental Management, 91*, 767–771.

Reddy, K. R. N., Salleh, B., Saad, B., Abbas, H. K., Abel, C. A., & Shier, W. T. (2010). An overview of mycotoxin contamination in foods and its implications for human health. *Toxin Reviews, 29*, 3–26.

Samuel, M. S., Sivaramakrishna, A., & Mehta, A. (2014). Degradation and detoxification of aflatoxin B1 by *Pseudomonas putida*. *International Biodeterioration and Biodegradation, 86*, 202–209.

Shim, W. B., Yang, Z. Y., Kim, J. S., Kim, J. Y., Kang, S., Gun-Jo, W., ... Chung, D. H. (2007). Development of immunochromatography strip-test using nanocolloidal gold-antibody probe for the rapid detection of aflatoxin B1 in grain and feed samples. *Journal of Microbiology and Biotechnology, 17*, 1629–1637.

Siegel, P. H. (2004). Terahertz technology in biology and medicine. *IEEE Transactions on Microwave Theory and Techniques, 52*, 2438–2447.

Stoik, C. D., Bohn, M. J., & Blackshire, J. L. (2008). Nondestructive evaluation of aircraft composites using transmissive terahertz time domain spectroscopy. *Optics Express, 16*, 17039–17051.

Stroka, J., & Anklam, E. (2002). New strategies for the screening and determination of aflatoxins and the detection of aflatoxin-producing moulds in food and feed. *Trends in Analytical Chemistry, 21*, 90–95.

Turner, N. W., Subrahmanyam, S., & Piletsky, S. A. (2009). Analytical methods for determination of mycotoxins: A review. *Analytica Chimica Acta, 632*, 168–180.

Ueno, Y., Rungsawang, R., Tomita, I., & Ajito, K. (2006). Quantitative measurements of amino acids by terahertz time-domain transmission spectroscopy. *Analytical Chemistry, 78*, 5424–5428.

Withayachumnankul, W., Fischer, B. M., Lin, H. Y., & Abbott, D. (2008). Uncertainty in terahertz time-domain spectroscopy measurement. *Journal of the Optical Society of America B, 25*, 1059–1072.

Xie, F., Lai, W. H., Saini, J., Shan, S., Cui, X., & Liu, D. F. (2014). Rapid pretreatment and detection of trace aflatoxin B-1 in traditional soybean sauce. *Food Chemistry, 150*, 99–105.