Raman spectroscopy combined with linear discriminant analysis for detection of pistachio Aflatoxins

Reza Mohammadigol [1]*, Mohammad Hadi Khoshtaghaza [2], Ali Mohammad Nikbakht [3], Bahareh Jamshidi [4], Majid Lashgari[1]

1- Assistant Professor, Biosystem Engineering Department, Faculty of Agriculture, Arak University, Arak, Iran.
2- Associate Professor, Agricultural Machinery Mechanics Engineering Department, Faculty of Agriculture, Tarbiat Modares University, Tehran, Iran.
3- Assistant Professor, Biosystem Engineering Department, Faculty of Agriculture, Urmia University, Urmia, Iran.
4- Assistant Professor, Agricultural Machinery Mechanics Engineering Department, Agricultural Engineering Research Institute, Karaj, Iran.
*Corresponding author, Email: r-mohammadigol@araku.ac.ir

Abstract

Aflatoxins contamination in Pistachio has been known as a serious problem for export market. Considering the difficulties of current experimental techniques for aflatoxins detection in pistachio (complication, time consuming and experience required), Raman spectroscopy combined with linear discriminant analysis technique was used to classify intact and aflatoxins contaminated pistachio samples (in concentrations of 100 and 20 ng/g). Raman spectra were analyzed in the spectral region of 848-2000 cm^{-1}. The effect of the preprocessing methods such as smoothing, normalization, the first and second derivatives, detrending, standard normal variate and multiplicative scatter correction and their combination on classification results was also studied. To evaluate the classification model, Sensitivity, harmonic mean of sensitivity and precision (F_β) of the classification results (confusion matrix) were calculated. Results indicated that Raman spectroscopy combined with linear discriminant analysis technique have potential to classify pistachio samples with and without aflatoxins. Moreover, all preprocessing methods reduced the average sensitivity and the average of F_β, except Savitzky-Golay (SG) smoothing method (F_β= 0.73, sensitivity= 72.22). The maximum value of F_β= 0.86 and sensitivity= 83.33 for intact samples classification using normalization or SG + normalization preprocessing methods were achieved.

Keywords: Aflatoxin; Pistachio; Preprocessing; Raman Spectroscopy

Introduction

Aflatoxins as the main family members of mycotoxins are carcinogenic metabolites which produced by species of Aspergillus, especially *A.flavus* and *A. parasiticus.* Aflatoxins are known as the most toxic mycotoxins because of their deleterious effects on human and animal health. So far, study on several aflatoxins have identified that B_1, B_2, G_1 and G_2 are the most common and important types of Aflatoxins (Arruse *et al.*, 2005). Maximum tolerated level of aflatoxin B_1 in pistachio is 12 ppb[1] in Europe and this amount varies in other countries from 2 to 20 ppb depends on their regulations. In some cases total aflatoxins B_1, B_2, G_1 and G_2 amount is considered as aflatoxin tolerated level (Anonymous, 2012). High Performance Liquid Chromatography (HPLC) with Immunoaffinity column method is mainly used for detection of aflatoxins in pistachio nuts and other agricultural products in the field of research and trade. Several published papers also confirm the accuracy

1- Particle per billion

of HPLC method (Dragacci *et al.*,2001 ; Pearson *et al.*, 1999 ; Stroka *et al.*, 2000). Aflatoxin extraction and cleaning up procedures using this method is complicated, time consuming and require the experience (Sinha, 1999).

An extensive research has been conducted on using Raman spectroscopy combined with different classification algorithms and methods for early detection and diagnosis of cancer tissues, bacteria, fungi and chemical drugs. The share of agriculture in this area is less than other areas such as chemistry, physics, biology, etc. However, it has been raising mainly due to the significant development on the optics tools that are required for Raman spectroscopy. Optical apparatus such as laser technology, detectors, microscopes, fiber optics and filters enhance the information obtained from Raman spectra of various materials (Slater *et al*., 2001). Classification or multivariate regression models for the spectral data interpretation have been used in different fields (Huang *et al*, 2010; The *et al*., 2008; The *et al*., 2010). The main advantage of Raman spectroscopy on biological materials is enhancement of information available on spectra (Das and Agrawal, 2011). Raman spectra can be seen as a "fingerprint" of inside chemical compounds, so if a toxic agent causes biochemical changes, the obtained Raman spectra will be changed (Notingher, 2007). Research reports have been indicated that the choice of an appropriate method for Raman spectra analysis is very role-playing to achieve the desired detection limit (Gryniewicz, 2012).

International trends in the production and supply of safe food has caused quick and reliable techniques for food quality and safe assessment have been demanded and the use of these technologies is dramatically growing. Regarding the difficulties of the conventional techniques and the potential of Raman spectroscopy, the main objective of this research is to detect pistachio aflatoxins using Raman spectroscopy.

Materials and methods

Samples preparation

Samples of pistachio (Var. Fandoghi) paste were prepared at three levels (Stroka *et al*., 2000): c- non-contaminated (15 spectra); a- 20 ppb aflatoxins contamination[2] (17 spectra); A- 100 ppb aflatoxins contamination[3] (18 spectra). Aflatoxins were purchased from Sigma Aldrich Ltd and diluted to the desired concentrations using deionized distilled water based on Beer-Lambert law (Parnis and Oldham, 2013).

Raman spectrometer

Raman spectra were collected using a SENTERRA dispersive Raman microscope (BRUKER Corporation, Germany) under a 785 nm diode laser excitation source over the range of 140–4000cm^{-1} for over 10 s exposure time (8 scans) at resolution of 9-15 cm^{-1}. Some pre-tests were conducted in order to achieve the best quality of Raman spectra signal and it was found that the 100 mW power has a higher signal to noise ratio. Therefore, the laser power was set to 100mW. The software package OPUS was used for spectra acquisition. Using 785 nm wavelength laser as the stimulating source through Raman spectroscopy process significantly causes decrease in florescent phenomenon in the spectra in addition to the appropriate sensitivity (Wang, 2013).

Preprocessing

A recorded Raman spectrum has unwanted data such as noise and background information. Therefore, spectral data preprocessing is necessary to achieve accurate and reliable calibration model. Preprocessing methods are various and each method has been developed for a specific purpose. Moreover, according to the different conditions of sampling, testing methods and tools, there is not possible to provide a specific method for all conditions and samples. Hence, conventional methods are usually evaluated in different spectroscopic studies to select the best method (Jamshidi, 2012). In this study, the effect of smoothing, normalization, first and second

2- 5 ppb of each type of aflatoxins ($B_1 + B_2 + G_1 + G_2$)
3- 25 ppb of each type of aflatoxins ($B_1 + B_2 + G_1 + G_2$)

derivatives, detrending, Standard Normal Variate (SNV) preprocessing methods, multiplicative scatter correction (MSC) and some of their combinations were evaluated.

Spectral data reduction

Typically, the number of samples (spectra) for successful training of machine learning classifier increases dramatically with increasing the number of variables (Raman shifts). Extracting principal components as a data reduction algorithm has been performed using principal component analysis (PCA) method in spectroscopic studies because of the existence of large number of spectral variables (Sigurdsson *et al.*, 2004 ; Wu *et al*., 2007). With the help of PCA a large number of variables are substituted with a limited number of new explanatory variables which are not correlated nor co-variant (Nicolai *et al.*, 2007).

Modeling

The potential of LDA based on the principal components (LDA-PCA) technique for Raman spectral classification was evaluated in this study. Desirable results obtained from the use of this technique in the classification of Raman spectra by other researchers (Nikbakht *et al*., 2011; Nikbakht, 2009) was the reason for choosing this technique. LDA maximizes the ratio of between class variance to the within class variance in each particular data set. Therefore, it guarantees maximal discrimination (Nikbakht, 2009). The Unscrambler (X10.3 Version) software was used for handling, modeling and analyzing the spectral data.

Model evaluation

To develop the models, 70% of the spectral data were selected randomly for training and the remaining 30% for testing (Jamshidi *et al*., 2012). After calibration, the model was fitted with test data set and the classification result was obtained in the form of a confusion matrix. This process was repeated identically for10 preprocessing methods. To evaluate the validity of classifying results and models generalizability, harmonic mean of sensitivity and precision measure (F_β) was used (Han *et al*., 2012) as is shown in Equation (1).

$$F_\beta = \frac{(1+\beta^2)\times Precision \times Sensitivity}{(\beta^2 \times Precision) + Sensitivity} \qquad (1)$$

$$(Sensitivity = \frac{TP}{P}, \quad Precision = \frac{TP}{TP+FP}$$

F_β is harmonic mean of sensitivity and precision, β is weight coefficient, TP is the number of true predictions, P is the number of samples, FP is the number of false predictions. To increase the evaluation precision, the weight coefficient (β) was determined as 2 (Han *et al*., 2012).

Safety considerations

Aflatoxins are highly toxic and are considered to be potentially carcinogenic. Safety clothing, nitrite gloves and safety glasses were applied. All materials in contact with aflatoxins were soaked in hypochlorite solution for deactivation of aflatoxins and all solvents and solutions were disposed of as hazardous waste.

Results and Discussion

A significant differentiation from photons scattering intensity was observed on 848-2000 cm-1 band (Figure1). In LDA technique, selecting the appropriate preprocessing method is based on trial and error. Therefore, before applying the data to the model, the common preprocessing was applied and their effects on the classification accuracy were compared (Tables 1 and 2).

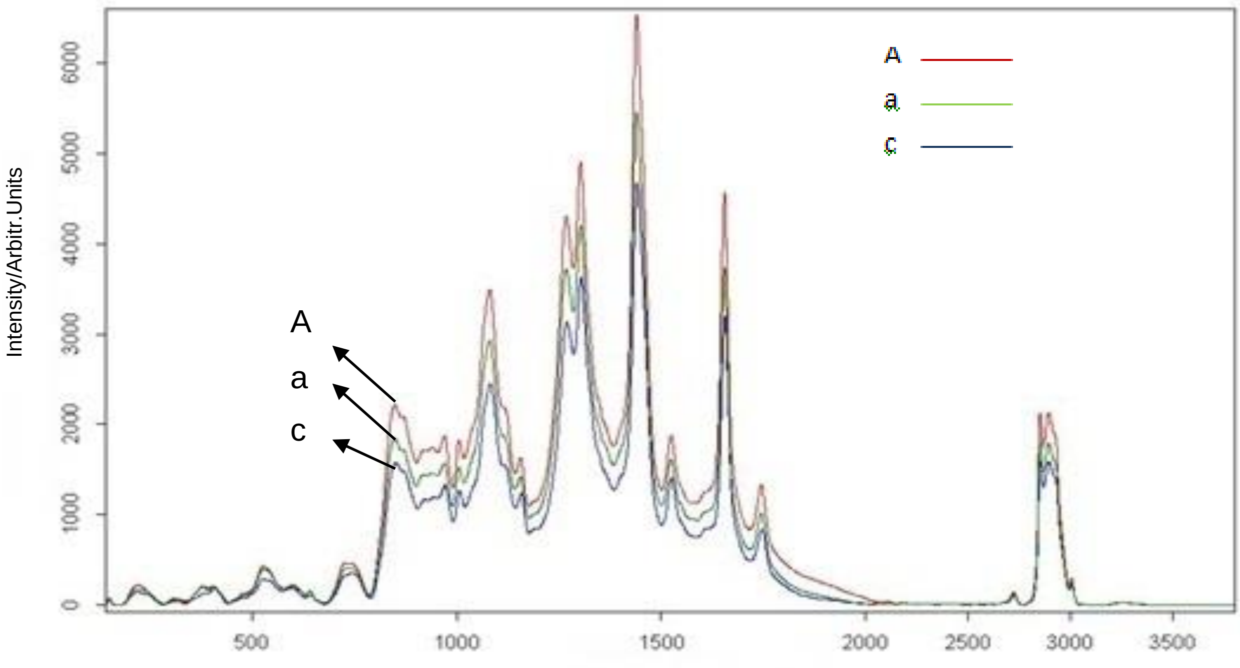

Figure 1. Full band mean Raman spectra of three classes.

In Tables 1 and 2, Fβ and sensitivity measures based on which preprocessing method has been done to the spectra, is presented. As can be seen, all studied preprocessing methods reduced the average sensitivity and the average F_β , except Savitzky-Golay (SG) smoothing preprocessing method (F_β= 0.73, sensitivity= 72.22). The maximum value of F_β= 0.86 and sensitivity= 83.33 for classification of aflatoxins-free samples were achieved using normalization or SG + normalization preprocessing methods.

Table 1. F_β measure due to different preprocessing methods (test data set)

Preprocessing	F_β			
	c	a	A	Average
Non- preprocessing	0.71	0.50	0.93	0.73
Normalization	**0.86**	0.28	0.89	0.70
Detrending	0.38	0.63	0.80	0.65
1stD+SG	0.00	0.65	0.00	0.35
2ndD+SG	0.00	0.65	0.00	0.38
SG+MSC	0.56	0.50	0.89	0.67
SG+SNV	0.56	0.50	0.89	0.67
SG	0.71	0.50	0.93	**0.73**
1stD+SG+Normalization	0.00	0.00	0.71	0.38
SG +Normalization	**0.86**	0.28	0.89	0.70

*:c- Non-Contaminated, a- 20 ppb contamination, A- 100 ppb contamination

Table 2. Sensitivity measure due to different preprocessing methods (test data set)

Preprocessing	Sensitivity (%)			
	c	a	A	Average
No preprocessing	66.67	50.00	100.00	72.22
Normalization	**83.33**	25.00	100.00	69.44
Detrending	33.33	75.00	80.00	62.78
1stD+SG	0.00	100.00	0.00	33.33
2ndD+SG	0.00	100.00	0.00	33.33
SG+MSC	50.00	50.00	100.00	66.67
SG+SNV	50.00	50.00	100.00	66.67
SG	66.67	50.00	100.00	**72.22**
1stD+SG+Normalization	0.00	0.00	100.00	33.33
SG +Normalize	**83.33**	25.00	100.00	69.44

As Tables 1 and 2 illustrate, F_β and sensitivity measures have a sharp drop in the1stD + SG, 2ndD + SG and 1stD + SG + Normalize preprocessing in all sample classes. Similar results was reported by other researchers (Ishikawa and Gulick, 2013). Primarily by increasing the derivative times, the Raman spectra peaks become more distinct, but also the existing noise is strengthened. Perhaps the LDA model weakness in category classification (based on applying 1st or 2nd derivative preprocessing) can be explained by this phenomenon.

Conclusion

The results of this research were proved the potential of Raman spectroscopy as a sensitive technique for Pistachio aflatoxins detection. Regarding to very low aflatoxins limit of detection (LOD) in pistachio, using Raman effect enhancing techniques to detect aflatoxins in the total form or individually on desired LOD need to further researches (such as enhancing substrates, classifiers and etc).

Acknowledgements

This research was supported by Tarbiat Modares University. The authors are grateful to Mrs. Hasannezhad and Mrs. Moghadasi for their help in spectra acquisition.

References

1. Anonymous, 2012. “Mycotoxins Legislation Worldwide,”.http://services.leatherheadfood.com/eman/FacSheet.aspx?ID=79.

2. Arrus. K, G. Blank, D. Abramson, R. Clear, and R. Holley, 2005. “Aflatoxin production by Aspergillus flavus in Brazil nuts,” *Journal of Stored Products Research*, Vol. 41, No. 5, pp. 513-527.

3. Das. R.S, and Y. Agrawal, 2011. “Raman spectroscopy: Recent advancements, techniques and applications,” *Vibrational Spectroscopy*, Vol. 57, No. 2, pp163-176.

4. Dragacci. S, F. Grosso, and J. Gilbert, 2001. “Immunoaffinity column cleanup with liquid chromatography for determination of aflatoxin M1 in liquid milk: collaborative study,” *Journal of AOAC International*, Vol. 84, No. 2, pp. 437-443.

5. Gryniewicz-Ruzicka. C. M, J. D. Rodriguez, S. Arzhantsev, L. F. Buhse, and J. F. Kauffman, 2012. “Libraries, classifiers, and quantifiers: A comparison of chemometric methods for the analysis of Raman spectra of contaminated pharmaceutical materials,” *Journal of Pharmaceutical and Biomedical Analysis*, Vol. 61, No. 0, pp 191-198.

6. Han. J, M. Kamber, and J. Pei, 2012. “Data mining: concepts and techniques,” Morgan kaufmann (An imprint of Elsevier),.

7. Huang. Z, S. K. Teh, W. Zheng, K. Lin, K. Y. Ho, M. Teh, and K. G. Yeoh, 2010. “In vivo detection of epithelial neoplasia in the stomach using image-guided Raman endoscopy,” *Biosensors and Bioelectronics*, Vol. 26, No. 2, pp. 383-389.

8. Ishikawa. S and V. Gulick, 2013, “An automated mineral classifier using Raman spectra,” *Computers & Geosciences*, Vol. 54, pp 259-268.

9. Jamshidi. B, 2012. “Non-destructive Taste Characterization and Classification of Oranges Using Vis/NIR Spectroscopy,” Ph.D. Thesis, Tarbiat Modares University (TMU), Tehran, Iran.

10. Jamshidi.B, S. Minaei, E. Mohajerani, and H. Ghassemian, 2012. “Reflectance Vis/NIR spectroscopy for nondestructive taste characterization of Valencia oranges,” *Computers and Electronics in Agriculture*, Vol. 85, pp 64-69.

11. Nicolai. B. M., K. Beullens, E. Bobelyn, A. Peirs, W. Saeys, K. I. Theron, and J. Lammertyn, 2007. “Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A revie,” *Postharvest Biology and Technology*, Vol. 46, No. 2, pp 99-118.

12. Nikbakht. A, 2009. “Model Determination for Intelligent and Non-Destructive Prediction of Qualitative properties of Tomato (one variety) using Raman Spectroscopy,” Ph.D. Thesis, Tarbiat Modares University (TMU), Tehran, Iran.

13. Nikbakht. A, T. Tavakkoli Hashjin, R. Malekfar, and B. Ghobadian, 2011. “Nondestructive determination of tomato fruit quality parameters using Raman spectroscopy,” *Journal of Agricultural Science and Technology*, Vol. 13, pp 517-526.

14. Notingher. I, 2007. “Raman spectroscopy cell-based biosensors,” *Sensors*, Vol. 7, No. 8, pp 1343-1358.

15. Parnis. J.M, and K.B. Oldham, 2013. “Beyond the Beer–Lambert law: The dependence of absorbance on time in photochemistry,” *Journal of Photochemistry and Photobiology A: Chemistry*, Vol. 26, No. 7, pp 6-10.

16. Pearson. S, A. Candlish, K. Aidoo, and J. Smith, 1999. “Determination of aflatoxin levels in pistachio and cashew nuts using immunoaffinity column clean-up with HPLC and fluorescence detection,” *Biotechnology techniques*, Vol. 13, No. 2, pp. 97-99.

17. Sigurdsson. S, P. A. Philipsen, L. K. Hansen, J. Larsen, M. Gniadecka, and H. C. Wulf, 2004. “Detection of skin cancer by classification of Raman spectra. Biomedical Engineering,” *IEEE Transactions*, Vol. 51, No. 10, pp1784-1793.

18. Sinha. K, 1999. “Testing methods for aflatoxins in foods,” *Food & Nutrition Bulletin*, Vol. 20, No. 4, pp. 458-464.

19. Slater. J.B, J.M. Tedesco, and R.C. Fairchild, 2001. “Raman Spectrometry and Its Adaptation to the lndustrial Environmen,” Handbook of Raman Spectroscopy: From the Research Laboratory to the Process Line, p. 41.

20. Stroka. J, E. Anklam, U. Jorissen, and J. Gilbert, 2000. “Immunoaffinity column cleanup with liquid chromatography using post-column bromination for determination of aflatoxins in peanut butter, pistachio paste, fig paste, and paprika powder: collaborative study,” *Journal of AOAC International*, Vol. 83, No. 2, pp. 320-340.

21. The. S. K, W. Zheng, K. Y. Ho, M. Teh, K. G. Yeoh, and Z. Huang, 2010. “Near-infrared Raman spectroscopy for optical diagnosis in the stomach: Identification of Helicobacter-pylori infection and intestinal metaplasia,” *International Journal of Cancer*, Vol. 126, No. 8, pp1920-1927.

22. The. S.K, W. Zheng, K. Y. Ho, M. Teh, K. G. Yeoh, and Z. Huang, 2008. “Diagnosis of gastric cancer using near-infrared Raman spectroscopy and classification and regression tree techniques,” *Journal of biomedical optics*, Vol. 13, No. 3, pp. 034013-034018.

23. Wang. Q, 2013. “Raman spectroscopic characterization and analysis of agricultural and biological systems,”. Iowa State University.

24. Wu. D, S. Feng, and Y. He, 2007. “Infrared spectroscopy technique for the nondestructive measurement of fat content in milk powder,” *Journal of dairy science*, Vol. 90, No. 8, pp 3613-3619.

نهمین کنگره ملی مهندسی ماشین‌های کشاورزی
(مکانیک بیوسیستم) و مکانیزاسیون
پردیس کشاورزی و منابع‌طبیعی دانشگاه تهران
۲ و ۳ اردیبهشت ۱۳۹٤ - کرج

استفاده از طیف سنجی رامان و تکنیک تحلیل شناسایی خطی برای تشخیص آفلاتوکسین پسته

چکیده

آلودگی پسته به آفلاتوکسین به‌عنوان معضل بازار صادرات این محصول شناخته شده است. با توجه به مشکلات روش‌های جاری تشخیص آفلاتوکسین‌ها در پسته (پیچیدگی، زمان‌بر بودن و نیازمند تجربه)، طیف‌سنجی رامان و تکنیک آنالیز شناسایی خطی برای طبقه بندی نمونه های سالم و آلوده به آفلاتوکسین پسته (غلظت های ۱۰۰ و ۲۰ نانوگرم در گرم) مورد استفاده قرار گرفت. داده‌های طیفی نوار cm^{-1} ۲۰۰۰–۸٤۸ طیف‌های رامان اخذ شده مورد تحلیل قرار گرفتند. در این پژوهش تأثیر روش‌های هموارسازی، نرمال سازی، مشتق اول و مشتق دوم، قوس گیری، توزیع نرمال استاندارد و تصحیح پراکنش افزاینده به صورت انفرادی یا ترکیبی در نتایج طبقه‌بندی مورد بررسی قرار گرفتند. به منظور ارزیابی مدل، حساسیت و شاخص میانگین هارمونیک دقت و حساسیت (ماتریس اغتشاش) نتایج طبقه‌بندی محاسبه شدند. نتایج نشان داد طیف‌سنجی رامان در ترکیب با تکنیک شناسایی خطی، پتانسیل طبقه بندی نمونه‌های پسته عاری و دارای آفلاتوکسین را دارد. اعمال پیش‌پردازش‌های مختلف به‌جز هموارسازی ساویتزکی گولای، باعث کاهش شاخص‌های متوسط میانگین هارمونیک ($F\beta$) و متوسط حساسیت مدل شدند (حساسیت= ۷۲/۲۲ و $F\beta$ =۰/۷۳). بیشینه مقدار $F\beta$ =۰/۸٦ و حساسیت= ۸۳/۳۳ طبقه‌بندی در جداسازی نمونه‌های سالم با اعمال پیش‌پردازش نرمال‌سازی یا هموارسازی+ نرمال سازی به دست آمدند.

کلمات کلیدی: آفلاتوکسین، پسته، پیش‌پردازش، طیف سنجی رامان