Application of Near Infrared Spectroscopy to detect fungal contamination in green coffee beans

Panchita Taradolsirithitikul[1] , Panmanas Sirisomboon[2], Cheewanun Dachoupakan Sirisomboon[1,*]

[1]*Department of Microbiology, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand*
[2]*Curriculum of Agricultural Engineering, Department of Mechanical Engineering, Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand*
**e-mail: cheewanun.d@chula.ac.th*

Abstract

Fungal contamination, especially mycotoxigenic fungal contamination in food can be harmful to human health due to their mycotoxin toxicity. The early detection of fungal contamination in raw materials at the beginning or during food processing can reduce the risk of the contamination from the food chain. Near Infrared Spectroscopy (NIRS) is considered as an alternative method to determine fungal and mycotoxin contamination in food. This research was to evaluate the possibility to use NIRS, with a wavelength range of 800 – 2500 nm, as a rapid method to detect fungal contamination in green coffee beans. The percentage of total fungal infection and NIR spectral data were obtained from 100 green coffee bean samples. Quantitative calibration models to detect the percentage of total fungal infection were developed using the original and pretreated absorbance spectra in conjunction with partial least square regression (PLSR). The best predictive model was developed from the multiplicative scatter correction pretreated spectra with the correlation coefficient (r) of 0.825, standard error of prediction (SEP) and bias of 16.730% and -0.129%, respectively. Results showed that NIRS technique had a great potential for the detection of fungal contamination in green coffee beans.

Keywords: Near Infrared Spectroscopy (NIRS), fungal contamination, PLS regression, green coffee beans

Introduction

Coffee is considered as the important economic plants for both of domestic consumption and exportation of Thailand. In 2011, the total export value of coffee beans, roasted coffee beans and coffee products are higher than 930 million baht (28 million USD) with more than 322,000 acres of growing area. There are two coffee varieties grown commercially for consumption: Robusta (*Coffee canephora*), and Arabica (*Coffee arabica*). Robusta coffee is the most highly production in the South, followed by Arabica coffee that is the main species produced in the North of Thailand (Department of Foreign Trade, Ministry of Commerce, 2012). Consumption demand is growing in Thailand and Thai people consume an average 1 g of coffee per day (FAOSTAT, 2009).

Coffee is one of the plants presenting a high ochratoxin A contamination. Ochratoxin A (OTA) is a mycotoxin, a secondary metabolite of filamentous fungi, produced by *Aspergillus* and *Penicillium*. OTA has well shown nephrotoxic, teratogenic and immunotoxic properties in animal (Varga and Kozakiewicz, 2006). The International Agency for Research on Cancer (IARC) classified OTA as a possible human carcinogen (group 2B) (IARC, 1993).
The using of an inappropriate farm management in pre or post-harvest can lead a high risk of fungal contamination in coffee beans, including, ochratoxigenic fungi, such as *Aspergillus*

section *Circumdati (A. westerdijkiae* and *A. ochraceus)* and *Aspergillus* section *Nigri (A. carbonarius* and *A. niger)* (Noonim et al., 2008). There are several studies reported the contamination of OTA in green coffee beans, roasted coffee beans and instant coffee (Noonim et al., 2008; Oscar et al., 2014). Because of the toxicity of OTA and its high occurrence in various kinds of food, the European Commission has fixed the maximum limits of OTA in the roasted coffee and the instant coffee at 5.0 and 10.0 µg/kg respectively (European Commission (EC) No. 1881/2006).

The early detection of ochratoxigenic fungal infection in postharvest or stored cereals and grains represents as a key opportunity to reduce the risk of OTA infection from the food chain. Traditionally, mycotoxigenic fungal infection in food has been detected using microbiological methods, including fungal enumerating using plate-counting or direct plating techniques, isolating in appropriate media and identifying to the genus and species level by morphological characterization such as macroscopic and microscopic characteristics (Pitt and Hocking, 2009). However, these methods are time consuming and requiring a well-equipped laboratory and highly skilled techniques. Also, there are difficulties regarding assay standardization and errors arising from contamination (Atkins and Clark, 2004). Currently, the polymerase chain reactions (PCR) method is considered a good alternative option for fungal molecular diagnostics due to high specificity and sensitivity (Niessen, 2007). However, PCR techniques have some limits due to protocol complexities, higher reagent costs and the choice of specific primers for each species (Santos et al., 2010).

Near Infrared Spectroscopy (NIRS) is a powerful technique for characterizing the chemical composition of the complex materials including food and agricultural products. It is a fast, nondestructive, environmental friendly and highly accurate method that requires little expert training. It has been used to analyze proteins in rice (Zhang et al., 2007), sugar content in fruits (Camps and Christen, 2009) and detected microorganisms in fruit juices (Al-Holy et al., 2006). This technique has been successfully applied to detect for mycotoxins and mycotoxigenic fungal contamination in agricultural products such as maize, wheat and rice. Pettersson and Aberg (2003) studied the use of NIRS for determination of the mycotoxin deoxynivalenol (DON) in cereals. Berardo et al. (2005) described the application of NIRS to assess maize, detecting the toxin contamination, fumonisin from *Fusarium verticillioides*. Singh et al. (2010) used a short wave NIRS hyperspectral imaging system in the range of 700 – 1100 nm to detect fungal contamination in wheat, including *Penicillium* spp., *Aspergillus glaucus* and *Aspergillus niger*. Dachoupakan-Sirisomboon et al. (2013) reported the potential application of NIRS with a wavelength range between 950 – 1650 nm for detection of fungi and aflatoxigenic fungi in rice.

The objective of this research was to evaluate the possibility of applying NIRS, with a wavelength range between 800 – 2500 nm for a detection of fungal contamination in green coffee beans.

Methodology

Coffee sampling

One hundred Arabica green coffee bean samples, harvested in the period of 2012 – 2013, were purchased from the different farms in the Northern part of Thailand.

Near Infrared scanning

Green coffee bean sample was placed in a quartz open cup and then scanned using NIRS, a Fourier Transform Near Infrared (FT-NIR) spectrometer (MPA, Bruker Optics, Germany) in a diffuse reflectance mode with over the range 12500 – 4000 cm^{-1} (800 – 2500 nm). Each sample was analyzed in duplicate (200 sub-samples). The scanning was done three times on each sub-sample and the total of 600 spectra were used for data analysis. After collecting spectra using the OPUS 7.0.122 software (Bruker Optics, Germany), all spectral data were exported into JCAMP-DX format in order to develop the chemometric analysis.

Mycological analysis

Direct plating method is considered one of the most effective methods for detecting, enumerating and isolating fungi from particulate food such as grains and nut (Samson et al., 2004; Pitt and Hocking, 2009). Twenty-five green coffee beans were randomly selected and directly plated on DG18 (Dichloran 18% glycerol agar) culture medium with 5 beans per plate. The plates were incubated at 25ºC for 5 – 7 days. After incubation, the number of infected beans was counted and the results were reported as the percentage of total fungal infection.

Model developing

The percentage of total fungal infection (reference data) was checked for outliers using Eq. (1) (Sirisomboon et al., 2012)

$$\frac{(x-\bar{x})}{SD} \geq 3 \qquad \text{Eq. (1)}$$

which is the Z score and when the Z score is ≥ 3, then the x value will be considered as outlier. x is the reference value of % total fungal infection, $\bar{x}$ is the average of % total fungal infection and SD is the standard deviation. All spectral data were checked for outliers by using principal component analysis (PCA) of the Unscrambler 9.8 software (Camo, Norway).

All NIRS models were developed by using the Unscrambler 9.8 software. The reference data were merged with spectral data and then separated into calibration set and prediction set in ratio of 7:3. Raw spectra and pretreated spectra in the calibration set were used to develop partial least square regression (PLSR) models. Spectral pretreatment was performed using multiplicative scatter correction (MSC), normalization (mean, maximum and range), first derivative according to Savizkye-Golay (11 and 21 wavelength points with polynomial order of 2), second derivative according to Savizkye-Golay (11 and 21 wavelength points with polynomial order of 2), standard normal variation (SNV), and baseline offset. Independent validation of the calibration models was achieved using the prediction set. The best model was selected using with the highest correlation (r), lowest standard error of prediction (SEP) and bias.

Results

Spectral data and reference data

Figure 1 shows the average of absorption NIR spectra of green coffee bean obtained over a wavelength of 800 – 2500 nm. In order to find the difference in the spectra between green coffee bean samples contaminated with 0 – 50% and 51 – 100% of total fungal infection, the average NIR spectra of green coffee beans were pretreated by applying the second derivative method of Savizkye – Golay (Figure 2). Both spectra display similar overall shapes with obvious peaks at 1420, 1703, 1726, 1759, 1909, 2051, 2266, 2307, 2344 and 2479 nm, respectively.

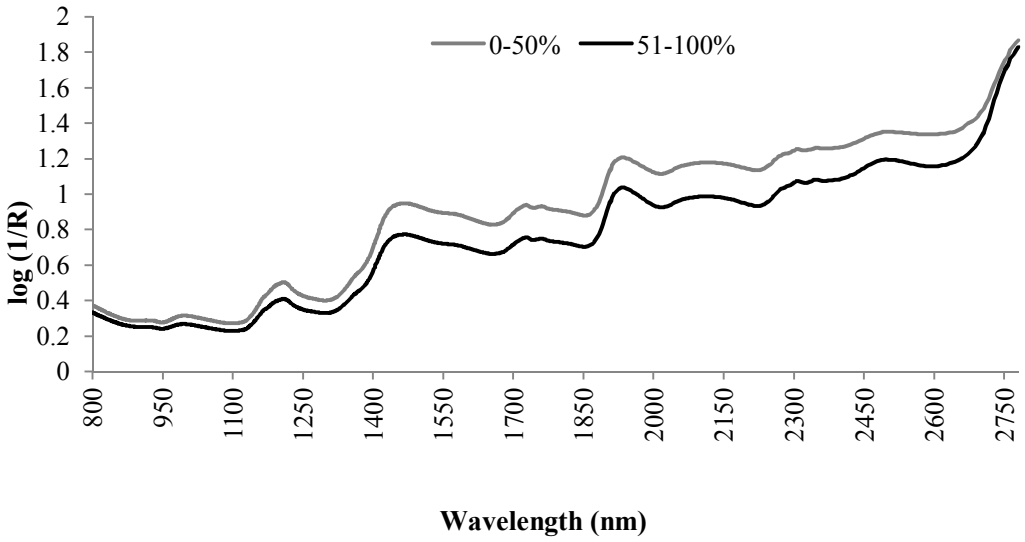
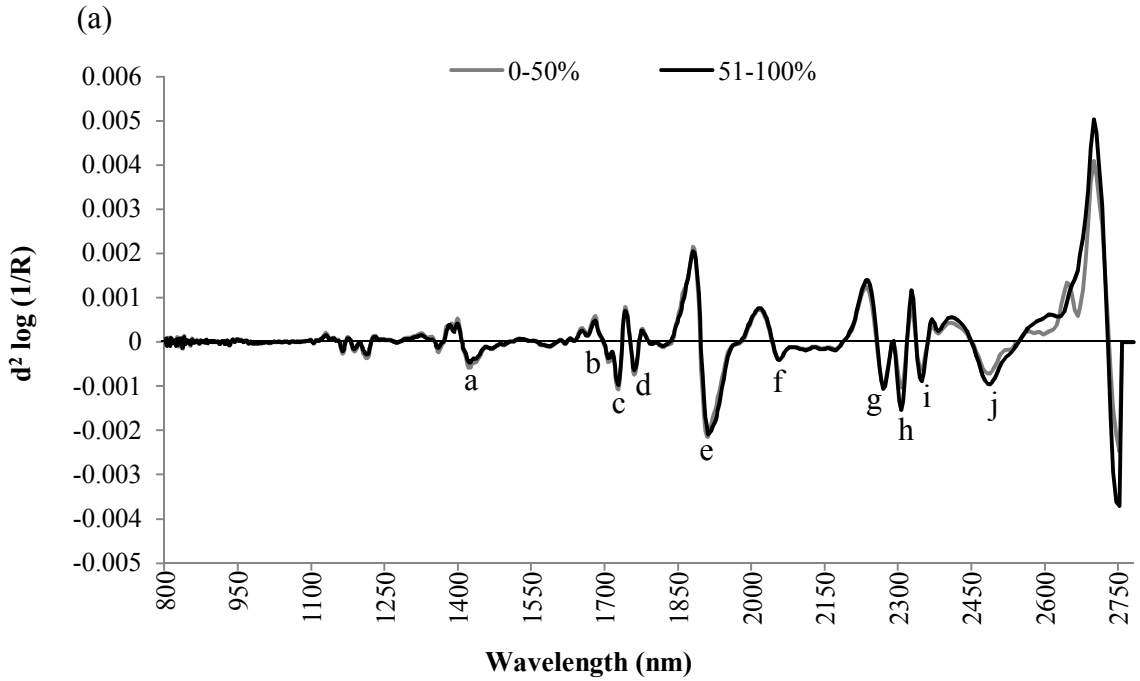


Figure 1: *Average raw NIR spectra of green coffee bean samples collected based on the difference percentage of total fungal infection.*



(b)

	Wavelength (nm)	Chemical bond
a	1420	the first overtone from the O-H stretching mode of ArOH
b	1703	the first overtone from the C-H stretching mode of CH_3
c	1726	the first overtone from the C-H stretching mode of CH_3
d	1759	the first overtone from the C-H stretching mode of CH_2
e	1909	the first overtone from the O-H stretching mode of POH
f	2051	the N-H symmetrical stretching mode of protein or N-H asymmetrical stretching mode of $CONH_2$
g	2266	the O-H stretching and O-H deformation of starch or O-H stretching and C-C stretching of starch
h	2307	the C-H stretching and C-H deformation of CH_2
i	2344	the CH_2 symmetrical stretching and $=CH_2$ deformation of $HC=CHCH_2$
j	2479	the C-H stretching and C-C stretching of starch

Figure 2: (a) Average pretreated NIR spectra by applying the second derivative method of Savizkye-Golay collected based on the difference percentage of total fungal infection and (b) chemical assignments of the observed NIR spectra (Osborne and Fearn, 1986).

The percentage of total fungal infection obtained from the direct plating method, ranged from 0 to 100%, with the mean of 83.56% and the standard deviation of 29.04%.

Data analysis and model developing

No outlier was present in both reference data and NIR spectral data. Then, all the data were separated into 2 sets, the calibration set and the prediction set. Table 1 shows the statistical values of the calibration set and the prediction set of the percentage of the total fungal infection model.

Table 1: Statistical values associated with the percentage of the total fungal infection in the calibration set and prediction set.

Set	n	Min	Max	Mean	SD
Calibration set	423	0	100	83.56	29.14
Prediction set	177	4	100	83.84	28.54

Table 2 shows the PLSR models developed from the raw NIR spectra and pretreated spectra. The model developed from raw spectra provided the highest correlation coefficient (r) of 0.839 with PCs, standard error of prediction (SEP) and bias of 11, 15.706% and -0.209%, respectively. For pretreated spectra, the model developed from multiplicative scatter correction (MSC) provided the highest correlation coefficient (r) of 0.825 with PCs, SEP and bias of 5, 16.730% and -0.129%, respectively. The RPD (ratio between the standard deviation (SD) of the reference method and the SEP) and RER (ratio between the range of the reference method and the SEP) values of the model developed from raw spectra and MSC pretreated spectra are 1.817, 6.112 and 1.706, 5.738, respectively. Figure 3 shows the correlations between the measured percentages of total fungal infection versus the predicted values from the NIR by PLSR model.

Table 2: PLS regression results for the percentage of total fungal infection model.

Pretreatment	PCs	Calibration			Prediction			RPD	RER
		R	SEC	Bias	R	SEP	Bias		
Raw spectra	11	0.821	16.622	2.697E-05	0.839	15.706	-0.209	1.817	6.112
Multiplicative scatter correction	**5**	**0.752**	**19.208**	**-3.152E-06**	**0.825**	**16.730**	**-0.129**	**1.706**	**5.738**
Mean normalization	7	0.769	18.622	-4.030E-07	0.800	17.342	-0.022	1.646	5.536
Max normalization	10	0.807	17.193	2.891E-05	0.828	16.167	0.146	1.766	5.938
Range normalization	10	0.808	17.150	-2.381E-05	0.830	16.113	0.148	1.772	5.958
1st derivative (11point)	11	0.858	14.988	-5.015E-07	0.791	17.477	-1.187	1.633	5.493
1st derivative (21point)	10	0.836	15.986	-2.024E-06	0.808	16.856	-0.285	1.693	5.695
2nd derivative (11point)	2	0.671	21.596	-6.000E-07	0.680	20.970	0.430	1.361	4.578
2nd derivative (21point)	6	0.767	18.698	-4.209E-07	0.779	17.954	-0.136	1.590	5.347
Baseline	7	0.782	18.154	-7.603E-06	0.800	17.336	0.909	1.647	5.538
Standard normal variation	4	0.753	19.172	-1.701E-06	0.778	18.130	0.545	1.574	5.295

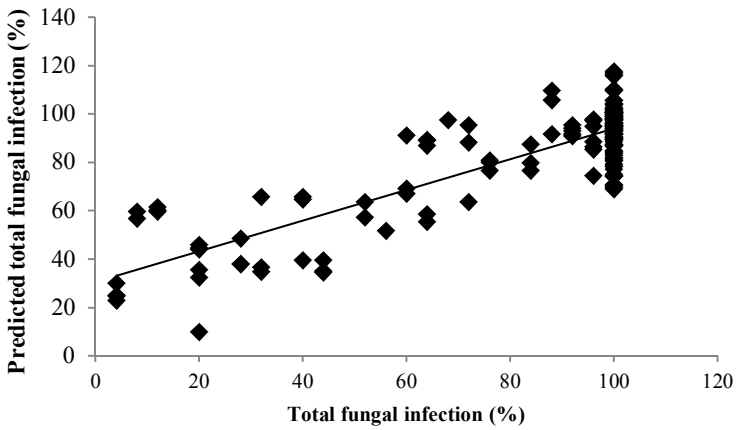


Figure 3: Correlation regression plot of the measured versus predicted the percentage of total fungal infection.

Figure 4 shows the regression coefficients obtained from the total fungal infection on PLSR model. This graph shows the spectral wavelengths that have a strong effect on the models such as at 1444, 1943 and 2488 nm.

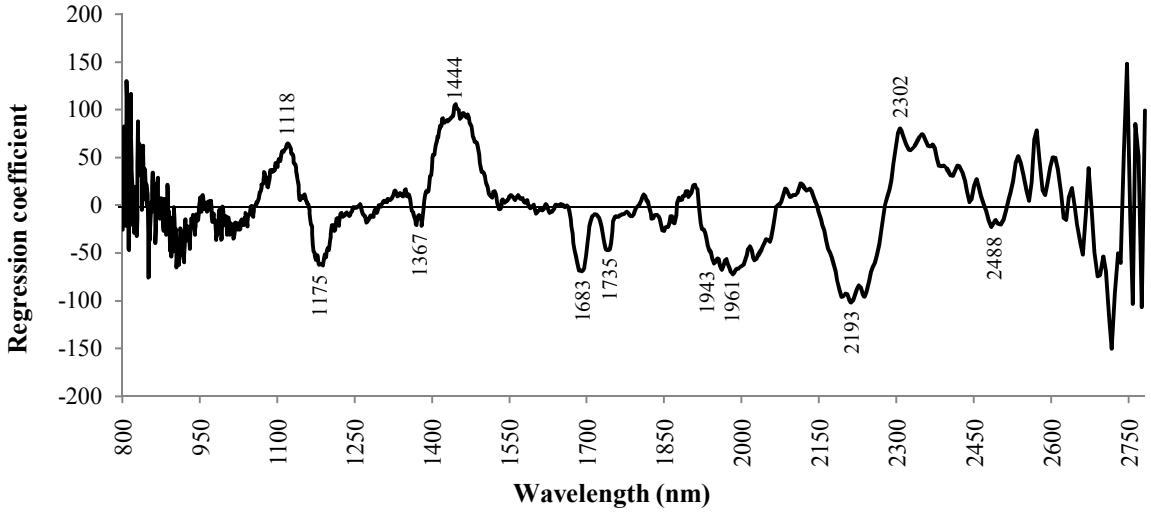


Figure 4: Regression coefficient plot for the percentage of the total fungal infection on PLSR model.

Discussion

Although the model developed from raw spectra provided the highest correlation coefficient (r) of 0.839 with 11 PCs, the best model of predicting total fungal infection in green coffee beans is the model that developed from multiplicative scatter correction (MSC) pretreated spectra (r = 0.825, PCs = 5). According to Villar et al. (2012), the model has a number of factors (PCs) higher than 6, the PLSR model were showed spectral noise and random errors and caused the overfitting. In MSC, the PLS factors used to develop the model was 5. This indicated that the model was more robust than the model from raw spectra. The value of r in this research indicated that the model can be used for screening and approximate calibration of total fungal contamination (Williams, 2007). In addition, the RPD and RER which indicated the performance of the model were 1.706 and 5.738, respectively. The value of RPD in this research (1.706) is in the range of 0.0 – 2.0, which showed this model is not suitable for classification and application (Williams, 2007). However, in the case of grains, flours and constituents, the RPD value as high as even 3.0 may be difficult to achieve, because of the complications with the sample preparation and sample presentation to the NIRS instrument (Williams, 2010 personal contact). Therefore, in this case, the lower values of RPD or RER gave a valuable indication of NIRS application.

Conclusion

The model developed showed good predictive performance. Results demonstrated the possibility to use NIRS technique as a rapid test for the detection of fungal contamination in green coffee beans. NIRS is easily and faster (2 – 3 min) comparing with conventional method used to detect fungal contamination in food. However, additional samples from different origins are required to improve the accuracy of the model.

Acknowledgements

This research was supported by The Asahi Glass Foundation. The authors would also like to thank The Bruker Corporation for providing the NIRS spectrometer (MPA) used in this research.

References

Al-Holy M., Lin M., Cavinato A. G. and Rasco B. A. (2006) The use of Fourier transform infrared spectroscopy to differentiate *Escherichia coli* O157:H7 from other bacteria inoculated into apple juice. Food Microbiol 23: 162-168.

Atkins S. D. and Clark I. M. (2004) Fungal molecular diagnostic: a mini review. J Appl Genet 45: 3-15.

Berardo N., Vincenza P., Paola B., Andrea S., Amedeo P. and Adriano M. (2005) Rapid detection of kernel rots and mycotoxins in maize by near-infrared reflectance spectroscopy. J Agr Food Chem 53: 8128-8134.

Camps C. and Christen D. (2009) Non-destructive assessment of apricot fruit quality by portable visible-near infrared spectroscopy. LWT-Food Sci Technol 42: 1125–1131.

Dachoupakan Sirisomboon C., Putthang R. and Sirisomboon P. (2013) Application of near infrared spectroscopy to detect aflatoxigenic fungal contamination in rice. Food Cont 33: 207-214.

Department of Foreign Trade, Ministry of Commerce. (2012) Coffee situation (Online) Source: http://www.dft.go.th (7 July 2014)

European Commission. (2006) Commission regulation (EC) No 1881/2006 of 19 December 2006, setting maximum levels for certain contaminants in foodstuffs. Official Journal of the European Union L 364: 1 – 25.

FAOSTAT. (2009) Crops Primary Equivalent (Online) Source: http://faostat.fao.org/site/609/DesktopDefault.aspx?PageID=609#ancor (2 January 2014)

IARC. (1993) Some naturally occurring substances: Food items and constituents, heterocyclic aromatic amines, and mycotoxins. IARC (International Agency for Research on Cancer) 56: 489.

Niessen L. (2007) PCR-based diagnosis and quantification of mycotoxin producing fungi. Int J Food Microbiol 119: 38-46.

Noonim P., Mahakarnchanakul W., Kristian F. N., Jens C. F. and Robert A. S. (2008) Isolation, identification and toxigenic potential of ochratoxin A-producing *Aspergillus* species from coffee beans grown in two regions of Thailand. Int J Food Microbiol 128: 197–202.

Osborne B. G., and Fearn T. (1986) Near infrared spectroscopy in food analysis. New York, Longman Scientific & Technical: 36-40.

Oscar G., Maritza A., Mario V. and Mario A. (2014) Occurrence of ochratoxin A in roasted and instant coffees in Chilean market. Food Control 46: 102-107.

Pettersson H. and Aberg L. (2003) Near infrared spectroscopy for determination of mycotoxins in cereals. Food Control 14: 229–232.

Pitt J. I. and Hocking A. D. (2009) Fungi and food spoilage (3rd ed.). New York: Springer. pp. 19-5, 275-295.

Samson, R. A., Hoekstra, E. S. and Frisvad, J. C. (2004) Introduction to food borne fungi. (7th ed.). Centraalbureau voor Schimmelcucutures, Netherlands.

Santos C., Fraga M. E., Kozakiewicz Z., and Lima N. (2010) Fourier transform infrared as a powerful technique for the identification and characterization of filamentous fungi and yeasts. Res Microbiol 161: 168-175.

Singh C. B., Jayasa D. S., Paliwala J. and White N. D.G. (2010) Identification of insect-damaged wheat kernels using short-wave near-infrared hyperspectral and digital colour imaging. Comput Electron Agr 73: 118–125.

Sirisomboon P., Tanaka M., Kojima T., and Williams P. (2012) Nondestructive estimation of maturity and textural properties on tomato 'Momotaro' by near infrared spectroscopy. J Food Eng 112: 218-226.

Varga J. and Kozakiewicz Z. (2006) Ochratoxin A in grapes and grape-derived products. Trends Food Sci Tech 17: 72–81.

Villar A., Gorritxategi E., Aranzabe E., Fernández S., Otaduy D. and Fernández L. A. (2012) Low-cost visible-near infrared sensor for on-line monitoring of fat and fatty acids content during the manufacturing process of the milk. Food Chem 135: 2756–2760.

Williams P. (2007) Near-Infrared Technology. Getting the Best Out of Light. Nanaimo B. C., Canada, PDK project Inc.

Zhang H., Wu J., Luo L., Li Y., Yang H., Yu X., Wang X., Chen L. and Mei H. (2007) Comparison of near infrared spectroscopy models for determining protein and amylose contents between calibration samples of recombinant inbred lines and conventional varieties of rice. Agr Sci China 6: 941-948.