DETECTION OF FUNGAL INFECTION IN PULSES USING NEAR INFRARED (NIR) HYPERSPECTRAL IMAGING

by

KANNAN KARUPPIAH

A Thesis Submitted to

the Faculty of Graduate Studies

in Partial Fulfillment of the Requirements

for the Degree of

MASTER OF SCIENCE

Department of Biosystems Engineering

University of Manitoba

Winnipeg, Manitoba

Abstract

Pulses are a major source of human protein intake nowadays and will continue to be so because of their high protein content. Pulse crops are members of the family *Leguminosae*. The five major pulse crops grown in Canada are chick peas, green peas, lentils, pinto bean and kidney beans. Over the past 20 years, Canada has emerged as the world's largest exporter of lentils and one of world's top five exporters of beans. These contribute more than $2 billion income to the Canadian economy. The major causes of fungal infection in these pulses are *Aspergillus flavus* and *Penicillium commune*. Early stages of fungal infections in pulses are not detectable with human eyes. Near infrared (NIR) hyperspectral imaging system is an advanced technique widely used for detection of insect infestation and fungal infection in cereal grains and oil seeds. A typical NIR instrument captures images across the electromagnetic spectrum at evenly spaced wavelengths from 700 to 2500 nm (a system at the University of Manitoba captures images in the 960 nm to 1700 nm range). From the captured images, the spatial relationships for different spectra in the neighborhood can be found allowing more elaborate spectral-spatial methods for a more accurate classification of the images. The primary objective of this study was to assess the feasibility of the NIR hyperspectral system to identify fungal infections in pulses. Hyperspectral images of healthy and fungal infected chick peas, green peas, lentils, pinto bean and kidney beans were acquired and features (statistical and histogram) were used to develop classification models to identify fungal infection caused by *Aspergillus flavus* and *Penicillium commune*. Images of healthy and fungal infected kernels were acquired at 2 week intervals (0, 2, 4, 6, 8 and 10 weeks from artificial inoculation).

Six-way (healthy vs the five different stages of infection) and two-way (healthy vs every stage of infection) models were developed and classifications were done using linear discriminant analysis (LDA) and quadratic discriminant analysis (QDA) classifiers. The LDA classifier identified with 90-94% accuracy while using the six-way model, and with 98-100% accuracy when using the two-way models for all five types of pulses and for both types of fungal infections. The QDA classifier also showed promising results as it identified 85-90% while using the six-way model and 96-100% when using the two-way models. Hence, hyperspectral imaging is a promising and non-destructive method for the rapid detection of fungal infections in pulses, which cannot be detected using human eyes.

Acknowledgements

First, I am very much thankful to Dr. Digvir S. Jayas, Vice President (Research and International) and Distinguished Professor, University of Manitoba for accepting me to do my research work at the University of Manitoba, Winnipeg, Canada and for being the supervisor of my Master's program. I am thankful for all the support, guidance, and help that he has given me throughout my program of study.

I offer sincere thanks to my committee member Dr. Noel D. G. White, Senior Research Scientist, Cereal Research Centre, Agriculture and Agri-food Canada for his continuous guidance and support throughout my research. I am thankful to him for his inputs to my research and for providing me with the pulse samples. He showed me new ways to do research.

I am thankful to Dr. Paul G. Fields, Senior Research Scientist, Cereal Research Centre, Agriculture and Agri-food Canada for his valuable suggestions during my research at the University of Manitoba. His positive attitude and supportive talks had immense influence on my research.

I specially thank Dr. Gabriel Thomas, Professor, Department of Electrical and Computer Engineering for being on my thesis committee and reviewing my progress during the program.

Let me say 'thanks' to R. Lankapalli, V. Chelladurai, and T. Senthilkumar for the support and suggestions they gave to me throughout my work. I also thank C.J. Demianyk of Agriculture and Agri-food Canada for his assistance in my research.

I thank my friends K.P. Kaja, E. Rajendren and C. Erkinbaev for their kind encouragement during my study.

I am greatly indebted to my parents, for giving me this wonderful life and educating me. Thanks to my sister and her family for their encouragement and great support in my life.

Finally, I acknowledge Department of Biosystems Engineering administrative and technical staff for providing support during my program.

Table of Contents

List of Tables

List of Figures

1. INTRODUCTION

Pulses belong to the family Leguminosae and are a major source of human nutrition because of their high protein content. Canada has become the world's largest exporter of pulses with more than $2 billion income to the Canadian economy. The major types of pulses grown in Canada are: chick peas (*Cicer arietinum* L.), green peas (*Pisum sativum* L.), lentils (*Lens culinaris* Medikus), pinto bean (*Phaseolus vulgaris pinto* L.) and kidney beans (*Phaseolus vulgaris* L.) with the major production coming from Saskatchewan which accounts for 79.3% of the total pulses in Canada (Statistics Canada, 2011). Pulses have high amounts of aspartic acid, glutamic acid, leucine, lysine, and arginine and also possess functional properties like water solubility, foaming and fat binding capacity (Joyce et al., 2010). Storage losses in pulses are a major problem because of fungal infection and insect infestation.

Fungal species especially *Aspergillus flavus* Link and *Penicillium spp.* can grow on many food commodities under suitable conditions and produce secondary metabolites, mycotoxins, which are toxic to humans and animals (Del Fiore et al., 2010). Food and Agricultural Organization (FAO, 2014) of the United Nations reported that more than 25% of world food production is affected by fungal infection. Fungus can damage a product, produce discoloration, release volatile compounds and release toxins. Hence, they need to be identified at the initial stages of infection so quick corrective actions can be implemented. There are many traditional methods available for identifying the toxigenic fungi but they are costly and time consuming. Hence, there is a need for a rapid, non-destructive, and accurate method for the detection of toxigenic fungi.

One such method is the use of hyperspectral imaging which allows the characterization of a sample's spectral (spectroscopic component) and spatial (imaging component) characteristics and is also a rapid and non-destructive method (Gowen et al., 2007). Hyperspectral imaging produces a hypercube, which can facilitate the analysis of extrinsic and intrinsic properties of samples. Several studies involving the application of hyperspectral imaging for quality assessment of food and agricultural products have been reported in the literature (Singh et al. 2007, 2009; Senthilkumar et al., 2012). Most of the research concerning this has been done in the Vis-NIR (visible near infrared) (400-1000 nm) or NIR (near infrared) (700-1700 nm) range and has been used to detect bruises in fruits and vegetables, fungal, and insect infestation in cereals and oilseeds.

For instance, hyperspectral imaging has been used to measure the moisture and oil content in maize (Cogdill et al., 2004), detection of insect damaged wheat seeds (Singh et al., 2009), detection of fungi in wheat (Singh et al., 2007) and detection of insect infestation in beans (Kaliramesh et al., 2013). There are no reported studies dealing with the detection of fungal infection in pulses using this technique.

Therefore, the objectives of this study were:

1. to develop an algorithm to classify healthy and fungal infected pulses;
2. to use this algorithm to identify fungal infection in five pulse types namely, chick peas, green peas, green lentils, kidney beans, and pinto beans; and
3. to compare the classification accuracies between different fungal infection stages (different weeks after inoculation).

2. LITERATURE REVIEW

2.1 Basics of hyperspectral imaging

Hyperspectral imaging (also known as imaging spectroscopy) as the name suggests, is an imaging technique, which collects both spatial and spectral information from the objects (Vadivambal and Jayas, 2015). Spectral information can be collected in any emissive region of the electromagnetic spectrum but visible and NIR spectra (Vis-NIR, 400 to 2500 nm) are commonly used for food quality assessment. The spectral information results from vibrational motion of molecules and vibrational-rotational modes through interactions with the electromagnetic spectrum and depends on chemical composition of the samples. Most published studies dealing with foods and agricultural materials have used 900 to 1700 nm wavelengths. The spectrum can identify and characterize a particular feature within the sensor's field of view because of unique spectral signatures or fingerprints that is a plot of reflectance or radiance versus wavelength at each pixel (spatial location).

2.1.1 Infrared Imaging

Infrared (IR) is invisible radiant energy, electromagnetic radiation with wavelengths longer than those of visible light ranging from 700 nm to 14000 nm (Favro et al., 2000).

2.1.2 Near-Infrared (NIR)

The NIR ranges from 800 nm to 2500 nm (Tamara et al., 2000), is defined by the water absorption, is used in fiber optic telecommunication because of low attenuation losses in the SiO_2 (silica) glass medium and also are commonly used in agriculture and soil analysis, raw material analysis, pharmaceuticals, moisture and biomass analysis and physical and chemical composition analysis.

2.1.3 Short-Wavelength infrared (SWIR)

In SWIR the wavelength ranges from 1550 nm to 1750 nm (Daoyi et al., 2005) and is characterized by increase in water absorption significantly at 1450 nm. The 1530 and 1560 nm range is the dominant spectral region for long-distance telecommunications.

2.1.4 Mid-Wavelength infrared (MWIR)

The MWIR ranges from 2500 nm to 6000 nm (Saleem et al., 2012) wavelength and is also referred to as the intermediate infrared (IIR). This is preferred when inspecting high temperature objects such as infrared thermography for civil and military applications where nondestructive evaluation and target signature identification are done.

2.1.5 Long-Wavelength infrared (LWIR)

The LWIR ranges from 8000 nm to 14000 nm (Saleem et al., 2012) wavelength and is used mainly in astronomy and space applications.

2.2 Applications of Hyperspectral Imaging in Agricultural and Food Sciences

Hyperspectral imaging is being used in all fields of agricultural and food sciences. A review of how hyperspectral imaging is used to assess cereals, pulses and oilseeds is discussed below.

2.2.1 Cereals

For the early detection of toxigenic fungi on maize Del Fiore et al. (2010) used hyperspectral imaging. They detected *Aspergillus niger* van Tieghem and *A. flavus*, on maize which can produce toxins that are poisonous to human beings. The main aim of their study was to discriminate between healthy and diseased maize kernels. The hyperspectral imaging system had an imaging based spectrometer ImSpector-Specim V10, working in the wavelength range

of 400-1000 nm on a desktop spectral scanner. The results predicted that this system was able to rapidly discriminate healthy and infected maize kernels 48 h after inoculation whereas traditional methods were not effective.

Mahesh et al. (2008) used hyperspectral imaging to differentiate Canadian wheat classes. They used Canada Prairie Spring Red (CPSR), Canada Western Red Winter (CWRW), Canada Western Soft White Spring (CWSWS), Canada Western Extra Strong (CWES), Canada Western Hard White Spring (CHWHS), Canada Western Red Spring (CWRS), Canada Prairie Spring White (CPSW) and Canada Western Amber Durum (CWAD) to assess the feasibility of hyperspectral imaging to differentiate between these different wheat classes. The hyperspectral imaging system consisted of an InGaAs NIR camera and scanned samples in wavelengths from 960-1700 nm at 10 nm intervals. The results showed that linear discriminant analysis (LDA) with a leave-one-out cross-validation method was able to accurately classify 100% for CPSR, CWRW and CWSWS wheat classes and greater than 94% for other wheat classes. The quadratic discriminant analysis (QDA) classifier with a leave-one-out cross-validation method was able to classify >86% for all the wheat classes.

Delwiche et al. (2011) used Vis-NIR hyperspectral imaging to detect *fusarium* fungal damage in wheat kernels. They imaged *fusarium* damaged wheat kernels in the wavelengths of 1000-1700 nm and used a LDA classifier. They used 144 and 125 wavelength pair images that correspond to NIR and visible systems, respectively, for the classification and were able to separate sound kernels from infected kernels with up to 95% classification accuracy. The NIR region produced better results than the visible region.

To discriminate variety and quality of rice, Wang et al. (2015) used hyperspectral imaging. Paddy rice samples were imaged at 400-1000 nm wavelengths and principal component

analysis (PCA) was used to select the region of interest. The back propagation neural network (BPNN) was used for the classification and the accuracy was up to 94.4%. The results demonstrated that this method could discriminate quality and variety of rice.

In another study for the detection of ergot bodies (sclerotia of *Claviceps purpurea* (Fr.) Tul.) in cereals, Vermeulen et al. (2012) used near infrared hyperspectral imaging. They collected images at 209 wavelengths (1100-2400 nm) using a NIR hyperspectral line scan camera and a conveyor belt. When wheat sample infected with 0.01% ergot bodies was imaged and analyzed using partial least squares discriminant analysis (PLSDA), it was able to easily detect and identify the ergot bodies and no error was found with non-contaminated wheat samples.

2.2.2 Pulses

Kaliramesh et al. (2013) used hyperspectral imaging to detect insect infestation in mung bean using images in the near-infrared region. They collected images in the wavelengths of 1000 nm to 1600 nm at 10 nm intervals and used the most significant wavelengths, which were 1100, 1290 and 1450 nm for classification. From each of these wavelengths, six statistical features namely maximum, minimum, mean, median, variance and standard deviation were extracted and used as inputs to the statistical LDA and QDA classifiers, which gave classification accuracies of 85% and 82% for identifying uninfested and infested mung bean kernels, respectively.

In another study, Kaliramesh et al. (2014) determined the main constituents in green gram using near-infrared hyperspectral imaging. Images of green gram were acquired at 960 nm to 1700 nm at 10 nm intervals and seventy five NIR reflectance intensities were extracted from each sample. They used principal components regression (PCR) and partial least squares

regression (PLSR) models for prediction. For moisture, protein, and starch content predictions of green gram, PLSR models demonstrated better performances than PCR models.

For predicting the sweetness and amino acid content in soybean, Monteiro et al. (2007) used a hyperspectral imaging system. Hyperspectral images of soybean were taken and artificial neural network based regression models were developed to calculate the level of fructose, glucose, sucrose and nitrogen contents and these results were compared with second derivative, raw reflectance and principal component analysis. The hyperspectral imaging method proved to be relatively accurate to predict the chemical content of soybean.

Huang et al. (2013) used hyperspectral imaging to detect insect damage in soybeans. They imaged soybeans in the wavelength region of 400-1000 nm and a total of four statistical features were extracted namely mean, median, maximum and standard deviation for classification. Classification accuracies were 100% using the support vector data description (SVDD).

2.2.3 Oilseeds

In order to detect fungal infection in canola, Senthilkumar et al. (2012) used near-infrared hyperspectral imaging. The canola seeds were artificially infected by *Aspergillus glaucus* group and images were acquired in the range of 1000 nm to 1600 nm at 60 evenly distributed wavelengths. Three significant wavelengths namely, 1100, 1230 and 1300 nm were identified using PCA and used in LDA and QDA classifiers for classifying the healthy and infected samples at second, fourth, sixth, eighth and tenth week from inoculation. Both LDA and QDA showed 99% accuracy for healthy samples and 100% for the fungal infected samples.

Hyperspectral imaging was used by Zhang et al. (2013) for detecting the distribution of nitrogen (N), potassium (K) and phosphorous (P)) and macronutrients content in oilseeds. They

imaged 140 leaf samples at 380-1030 nm wavelength range and partial least square regression (PLSR) and least-squares support vector machines (LS-SVM) were used for the analysis and classification. The regression coefficients for the LS-SVM model was 0.882 and for PLSR was 0.746, which best predicted the three macronutrients present in oilseeds.

In another study, Zhang and He (2013) used hyperspectral imaging to predict rapidly the estimation of seed yield in oilseed rape. Images were taken in the wavelength range of 380-1030 nm range and PLSR was used for the analysis. The wavelengths 543, 686, 718, 741, 824 and 994 nm produced good results with a root mean square error value of 23.72. These wavelengths best predicted seed weights of individual plants and the seed yield.

2.2 Applications in Fruits and Vegetables

2.2.1 Fruits

The potential of hyperspectral imaging to detect bruises in apples was studied by ElMasry et al. (2008) wherein they used the spectral region between 400 and 1000 nm. For selecting the dominant wavelengths partial least squares method and discrimination analysis were used and 750, 820 and 960 nm wavelengths were selected for multispectral imaging tests. The classification results predicted that this system could detect apples easily after 1 h of bruising and it can successfully separate bruised apples from sound apples.

Xing et al. (2005) used hyperspectral imaging to detect bruises on Jonagold apples. They imaged a total of 160 apples in the wavelength range of 400-1000 nm. Principal component analysis (PCA) was used to select efficient wavebands and they developed algorithms for bruise and stem/calyx detection. The classification accuracy of 1 day bruises was 84.6 and 77.5%, respectively and none of the sound apples were misclassified as stem or calyx.

In order to identify the tissue characteristics of different types of cherries, a machine vision system was created by Guyer and Yang (2000). It consisted of an enhanced NIR range Vidicon black and white camera (range 400-2000 nm), a computer and a monochrometer controlled light source. The spectral images of multiple cherry samples were taken over 680-1280 nm range at 40 nm intervals. A multilayer feed-forward artificial neural network was used for pixel-wise classification. About 73% classification accuracy was achieved for correct identification and quantification of all types of defects in cherries.

To detect fungal infection in date fruits, Teena et al. (2014) used NIR hyperspectral imaging. They used three groups of samples namely untreated control (UC), sterile control (SC) and inoculated samples (IS) which were artificially infected with *A. flavus*. Images were taken at 960-1700 nm wavelengths at 10 nm intervals at every two days after inoculation for ten days. Top four wavelengths were selected by using PCA and a total of 64 features were extracted and used for classification using LDA and QDA. For IS, the classification accuracy (control vs each infection stage) was 91.5%, 91% and 99% and it was 92.4%, 100% and 99.6% for UC. The QDA classifier yielded better classification results than the LDA.

2.2.2 Vegetables

For the identification of bruises on pickling cucumbers, Ariana et al. (2006) used a near infrared hyperspectral imaging system. They imaged pickling cucumbers at 0-3 days and 6 days in the wavelengths of 900-1700 nm after they were mechanically damaged by rolling load. PCA was used and the wavelength region 950-1350 nm was chosen for classification. The accuracies ranged from 75 to 95% for the first 6 days after bruising and the analysis revealed that general classification performance analysis was better than the PCA.

Siripatrawan et al. (2011) used hyperspectral imaging to rapidly detect *Escherichia coli* (Migula), *Casteellani and Chalmers* contamination in freshly packed spinach. Hyperspectral images were taken in 400-1000 nm wavelength range and principal component analysis and artificial neural network were used for the analysis and classification of data, respectively. A prediction map was created for the classification and this method proved to be a rapid one for detecting *E. coli* contamination in fresh spinach.

2.2.3 Quality Measurement of Fruits and Vegetables

Quality evaluation factors such as nutritive values, chemical constituents, sensory attributes, functional and mechanical properties of fruits and vegetables were studied by Abbott (1999). He used a hyperspectral imaging system to detect the following properties: mechanical properties related to texture, chemical properties related to taste and aroma, and optical properties related to appearance. The imaging system was more accurate than human inspection and it can serve as a vital tool for research and inspection.

2.3 Application in Meat and Meat Products

Kamruzzaman et al. (2011) used NIR hyperspectral imaging system to discriminate three types of lamb muscles. Samples from *Psoas Major* (PM), *Longissimus dorsi* (LD) and *semitendinosus* (ST) of Charollais breed were imaged in the wavelengths of 900-1700 nm. PCA was performed for selecting wavelengths (934, 974, 1074, 1141, 1211 and 1308 nm) and reducing dimensionality of data. The results showed 100% classification accuracy for all the three types of lamb muscles.

For poultry surface fecal contaminant detection, Park et al. (2006) used hyperspectral imaging system with four preprocessing methods on 64 birds. Using 2-wavelengths (I_{565}/I_{517}) and 3-wavelengths $(I_{576} - I_{616})/(I_{529} - I_{616})$ equations, a band-ratio image processing algorithm was applied. This method was able to predict surface contaminants with as high as 96.2% accuracy.

The potential of hyperspectral imaging system to assess pork quality based on texture, color and exudation characteristics was studied by Qiao et al. (2007). The pork samples were imaged and PCA was used to compress the data into 5, 10 and 20 principal components and artificial neural network (ANN) was used for the classification. The feed-forward neural network yielded 69% and 85% accuracies for 5 and 10 PCs, respectively.

2.4 Insect Infestation Detection

The NIR is a nondestructive tool for detection of insect infestation in pulses, particularly Mung bean (*Vigna radiata* (L.) R. Wilczek), which is one of the major pulse crops grown in India. Beans were artificially infected with Cowpea weevil (*Callosobruchus maculatus* (F.)) (Kaliramesh et al., 2013). The NIR hyperspectral imaging system was used and images were acquired in the wavelengths of 1000-1600 nm at 10 nm intervals and three significant wavelengths were identified as 1100, 1290 and 1450 nm from which statistical features such as minimum, maximum, mean, median, standard deviation, variance and ten histogram features were extracted and used in the statistical classifiers. These classifiers predicted accuracies of more than 85% and 82% for the uninfected and infected pulses, respectively. Mung bean in pupal and adult stages had higher classification accuracies than in the egg and larval stages.

Canadian wheat gets rejected because of the presence of storage insects, especially *Sitophilus oryzae* (L.), *Rhyzopertha dominica* (F.), *Cryptolestes ferrugineus* (Stephens), and *Tribolium castaneum* (Herbst). Infestation levels by insects were studied by Singh et al. (2009). The wheat seeds were artificially infested with these insects and images were scanned in the wavelengths of 1000-1600 nm using the NIR hyperspectral imaging system and the significant wavelengths were selected using the multivariate image analysis. Six statistical image features (maximum, minimum, mean, median, standard deviation and variance) and 10 histogram features were extracted from the images at 1101 nm and 1305 nm and were used in statistical discriminant classifiers (linear, quadratic). Linear and quadratic discriminant analysis classifiers classified 85-100% of healthy and insect damaged wheat seeds.

2.5 Fungal Detection

Del Fiore et al. (2010) detected the toxigenic fungi on maize using the SWIR hyperspectral imaging system. Fungi can grow on many food commodities and their secondary metabolites are toxic for humans and animals. The maize seeds were artificially infected with *Aspergillus flavus*, *A. parasiticus* (Speare) and *A. niger* and imaged using a hyperspectral imaging system. A desktop computer based imaging spectrometer (ImSpector-Specim V10) was used in the wavelengths of 400-1000 nm. After 48 h of inoculation this system was able to rapidly identify the maize seeds infected with toxigenic fungi whereas traditional methods were not able to complete this task.

The potential of NIR hyperspectral imaging system to identify fungal infection of wheat on the field and in storage facilities was studied by Singh et al. (2007). Wheat seeds were

artificially infected with *A. glaucus* and *A. niger* and images were taken in the wavelengths of 1000 to 1600 nm. Multivariate image analysis was used to reduce the data size and two-class and four-class models were developed by k-means clustering discriminant (linear, quadratic and Mahalanobis) analyses. The two-class models were able to classify 100% for healthy seeds and 97.8% for infected seeds and the four-class models were able to accurately identify 91.7% for healthy seeds and 95% for the infected seeds.

2.6 Application of hyperspectral imaging in Health Sciences

Khoobehi et al. (2004) used hyperspectral imaging to monitor the relative spatial changes in retinal oxygen saturation. Images were taken under pure oxygen and room air at 15 mm Hg (normal) and 60 mm Hg (5 min) and saturation maps of oxygenated and deoxygenated hemoglobin were extracted. The results obtained from hyperspectral imaging proved to be better when compared with experimental outcomes.

In a clinical environment for noninvasive and quantitative analysis of human tissue, hyperspectral imaging was used in visible reflectance mode (Zuzak et al., 2002). The imaging system was tuned to take images in the wavelengths of 525-725 nm using the liquid crystal tunable filter (LCTF). The hyperspectral image cube was quantified in terms of apparent absorbance and compared with that of the clinical model obtained by using oxyhemoglobin and deoxyhemoglobin signals in the 525-645 nm range.

For identifying the functional maps of the retina, Johnson et al. (2007) developed a retinal imaging spectroscopy using chromophore spectra. A snapshot imaging spectrometer obtaining both spectral and spatial data in the range of 450 to 700 nm with 50 bands each was used which

eliminates pixel mis-registration and motion artifacts as compared with normal imaging. From this and using three-wavelength algorithm, oxygen saturation maps were obtained which predicted 95% for the healthy subject arteries and 30 to 35% for the veins, respectively.

2.7 Fungal Detection Methods

2.7.1 Molecular and Microscopic methods

For detecting airborne fungal spores, Williams et al. (2001) used PCR based and integrated air sampling methods. Their spores were collected using a miniature cyclone-type air sampler and were counted using a microscope after suspending in 0.1% Nonidet P-40. They tested three specific assays namely: PCR, single-step PCR and nested PCR followed by southern blotting and probing. Detection sensitivity reduced when there were high concentrations of pollen, dust and unidentified fungi. Out of all the DNA extraction methods, the nested PCR easily detected fungal spores as compared to the single-step PCR or PCR.

Einsele et al. (1997) developed a PCR assay for the identification and detection of *Aspergillus* and *Candida* species using 134 fungal and 85 non-fungal isolates. The results were also verified using clinical trials wherein they divided patients into four groups: Group A with a total of 35 patients were kept as controls, Group B to D with a total of 86 patients in each group with febrile neutropenia and without fungal colonization. Out of these none of the A group patients and only 3 from the groups B to D were tested PCR positive. The preliminary data provided insights for early detection and monitoring of antifungal therapy.

2.8 Summary

There is a place for using near infrared hyperspectral imaging system to identify the fungal infection in pulses. Biochemical changes are observed when the stored pulses are affected by fungal infection and they cannot be seen with a normal camera in the visible range or by human eyes. Fungi also consume chemical components for their growth, hence there are only subtle differences between low and highly infected pulses. A lot of work has been done for the insect identification in pulses (Kaliramesh et al., 2013) but for fungal infection not many studies have been done. Thus, there is a need to assess the ability of near infrared hyperspectral imaging system to detect fungi in stored pulses.

3. Materials and Methods

3.1 Fungal Sample Preparation

The fungal species, namely *Penicillium commune* Thom, C. and *Aspergillus flavus*, were acquired from the Cereal Research Centre and Food Science Department, University of Manitoba, Winnipeg, Canada. The fungal species were made to grow in large numbers by placing infected seeds on filter paper saturated with 7.5 mL aqueous NaCl solution in petri dishes. Seven days later, pure fungal lines were placed for one week at 30°C on potato dextrose agar. These agar moulds were then placed in 200 mL sterilized water and shaken in a plastic spray bottle mixed with 1 drop of Tween 20. The five pulse types namely chick peas, green peas, green lentils, pinto bean and kidney beans were obtained from the Saskatchewan Pulse Growers, Saskatoon, Canada. Each of the pulse types were moisturized to 17% moisture content (wet basis) by adding predetermined quantities of distilled water and mixing in a rotating drum. The moisturized pulses were surface sterilized by soaking in 1% sodium hypochlorite for 2 min. They were then rinsed thoroughly using sterilized water and dried on a paper towel for 2 h. Seeds were then sprayed with fungus-inoculated water until all seeds were wet. The control samples were maintained at a moisture content of 15% (wet basis, no inoculation).

3.2 Hyperspectral Imaging System

3.2.1 Hardware

The system used for this research is the same as described by Senthilkumar et al. (2012) and Kaliramesh et al. (2013) and is briefly described here. A short wave near infrared hyperspectral imaging system (SWNIR) consisted of a FFT-CCD area-scan image sensor (Model No. C7042, Hamamatsu Photonics, Hamamatsu, Japan) working in 900-1700 nm range with

image size 640 × 480 pixels. The camera sensor was made up of indium gallium arsenide (InGaAs). Two liquid crystal tunable filters (LCTFs), which allowed wavelengths only at specified value, were mounted on the camera and offered excellent wavelength selection and imaging quality.

Data acquisition was done using the PCI data acquisition card (National Instruments PCI-1422., Austin, TX). The card could acquire images from a wide range of digital cameras. The image acquisition card was compatible with the RS-422 signals provided by the InGaAs camera and the camera was mounted on a copy stand. This stand contained a pair of 300 W halogen lamps (one on each side of the camera system, mounted at 45° angle). Halogen lamps emit light in the visible-near infrared (400-2500 nm) region of the electromagnetic spectrum suitable for spectroscopic imaging applications (Ushio Lighting Inc., Cypress, CA).

The whole system (Figure 3.1) was connected by data cables that served two purposes: the first to provide and receive vital signals from the camera and second to provide all the DC power requirements. The LCTF cable linked the LCTF to the LCTF controller and the COM cable linked the PC to the LCTF controllers.

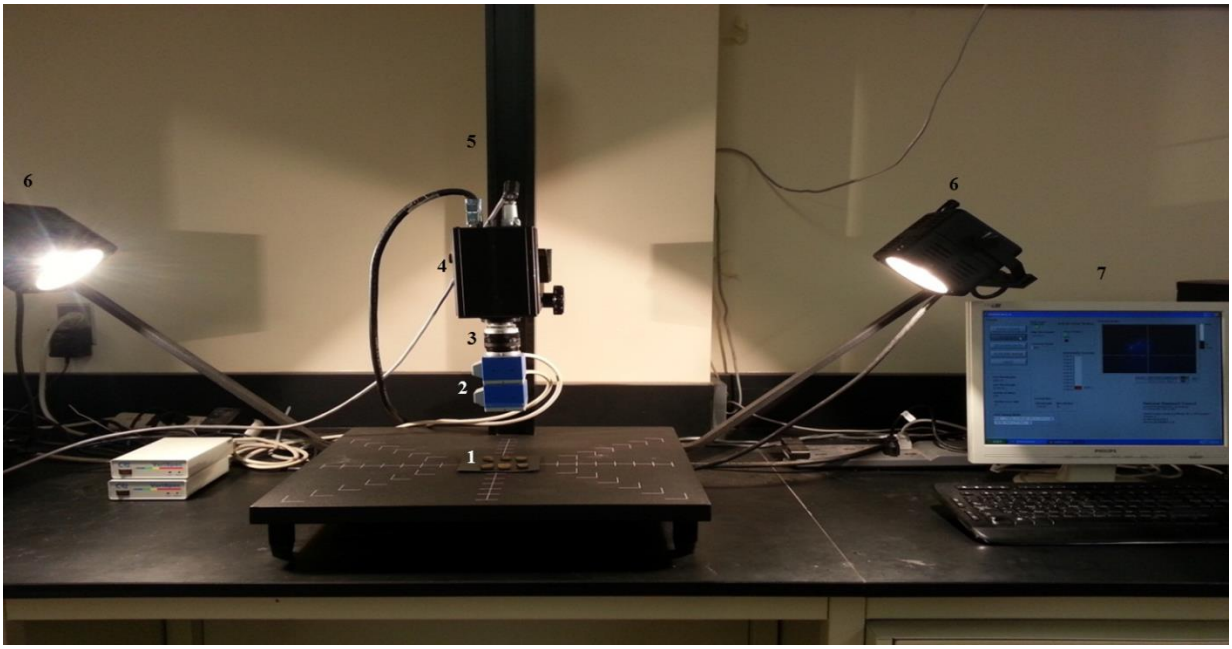

Fig. 3.1 Short-wave near-infrared hyperspectral imaging system 1. Pulse sample, 2. Liquid crystal tunable filter (LCTF), 3. Lens, 4. NIR camera, 5. Copy stand, 6. Illumination (Halogen-tungsten lamp), 7. Data processing system

3.2.2 Software

The software provided a user interface, which had different modes of operation. The mode of operation was selected by the user and was shown in the current mode indicator. There was an option for co-adds which must be picked prior to collection of spectroscopic data. After this the dark count image was collected which represented the inherent noise that was subtracted from the collected spectroscopic data. The dark count indicator was turned ON once the dark count was collected; it remained constant for that imaging session.

Then the user had the option to align the data using the alignment mode. During this mode, the wavelength of the LCTF filters was changed to the desired value. Then the data were acquired using the Acquire data mode from the first wavelength to the last wavelength where the

number of slices defined the number images to be collected. This mode collected a single spectroscopic image for the full range of wavelength (960-1700 nm).

Upon startup, the initial status of LCTF filters was examined. If LCTFs were not initialized the software would not proceed to the next stage. A blinking light on the LCTF filters indicated that the filters were initializing and once this was done the filters initialized indicator changed to green indicating that the hardware was ready to collect the data.

There were two modes for the LCTF filters for the user to choose from: in the first mode, the wavelength can be changed without confirming and in the second mode the wavelength that was changed must be confirmed using the check LCTF status option. For greatest LCTF reliability during data acquisition the user must choose mode 2 but the acquisition rate becomes slower than that of mode 1. In this study mode 1 was used.

Using the data acquisition and alignment mode the 640 × 480 pixel image slice was displayed. Scaling of the images can be done manually using a bar graph. This gave an option to the user to move it to a desired setting. It was assigned a color value of white and least intense pixel value was assigned a color value of black. All intermediate values were assigned various shades of blue. After this, data were acquired and stored as a .spc file, which was used later for processing.

3.2.3 Image Acquisition

Three hundred kernels of each of the five (chick peas, green peas, lentils, pinto bean and kidney beans) pulse types at second, fourth, sixth, eighth and tenth week from inoculation and healthy samples were imaged. The dark and white count (99% reflectance) was taken at regular

intervals. Six randomly selected seeds for each sample were manually placed on a black plastic board without touching each other in the field of view (FOV) of the NIR hyperspectral imaging system (Figure 3.1). The images were acquired using a Lab VIEW programme (Version 7.1, National Instruments Corp., Austin, TX). The images were taken in the usable wavelength region of 960-1700 nm and the camera was aligned to the center wavelength of 1300 nm. A total of 1800 images of individual kernels of each pulse type were imaged and used for further analysis.

3.2.4 Data Analysis

The image files acquired were imported to MATLAB (Mathworks Inc., Natick, MA, USA) for displaying and analyzing the hyperspectral data using a MATLAB programme. The signal-to-noise ratio (SNR) was improved by co-adding 10 images at each of 75 wavebands. Due to sensor defects (these are normal for any camera) the images had a few dead pixels which were removed using a 3X3 median filter. The original hypercube was changed to grey scale image intensity format, then using thresholding to binary image and all the seeds were labeled using the connected pixels.

The six labeled seeds were analyzed individually using a multivariate image analysis (MVI) programme written in MATLAB in which principal component analysis (PCA) was performed (Geladi and Grahn, 1996). The hypercube was a three-dimensional array because the image had width, height and a wavelength value as variable index and it was reshaped to pixel intensities by using reflectance values for all the 75 wavelength values. Hence, a k x 75 sized two dimensional arrays was obtained (k – the total number of pixels at each waveband for the labelled seeds). Finally, the PCA was applied to this two-dimensional data set for each pulse

type. Three highest factor loadings from the first principal component and two highest factor loadings from the second principal component were selected as the most significant wavelengths. From these significant wavelengths for each sample, statistical features such as maximum, minimum, mean, median, variance, standard deviation and 10 histogram features for each wavelength (a total of 50) were calculated.

3.3 Classification and Model Development

For the classification of different stages of infections of five pulse types by two fungal species and healthy samples, linear discriminant analysis (LDA) and quadratic discriminant analysis (QDA) were used.

3.3.1 Linear Discriminant Analysis (LDA)

The LDA is a method used in statistics, machine learning and pattern recognition to find linear combination of events and separates two or more classes as a generalization of Fishers linear discriminant and is primarily used for data classification and dimensionality reduction. This method of classification is useful where the within class frequencies are unequal and random test data are used.

Two different approaches to LDA are class-dependent transformation (maximizing the ratio of between class variance to within class variance) and class-independent transformation (maximizing the ratio of overall variance to within class variance).

The following are the steps involved in LDA:

1. Formulating the data set and the test sets, which are to be classified.

   Set 1= $[a_{11}\ a_{12}$ Set 1= $[b_{11}\ b_{12}$

   $a_{21}\ a_{22}$ $b_{21}\ b_{22}$

   … … … …

   … … … …

   $a_{m1}\ a_{m2}]$ $b_{m1}\ b_{m2}]$ ……. (1)

2. Computing the mean of entire data set and each data set. Let μ_1 and μ_2 be the mean of data set 1 and 2, respectively and μ_3 the mean of entire data set, then,

   $$\mu_3 = p_1 \times \mu_1 + p_2 \times \mu_2 \quad \ldots\ldots. (2)$$

   where p_1 and p_2 are the apriori probabilities of the classes.

3. Formulating the criteria for separability for between-class and within class by means of a covariance matrix.
4. Calculating the Eigen vector of the transformation which represents the 1-D invariant subspace of the vector space. Linear dependency between features is indicated by a zero Eigen value.
5. Transforming the Eigen values using LDA transforms and computing the Euclidean distance used to classify data points.
6. The smallest Euclidean distance classifies the test vector among n distances as belonging to class n.

3.3.2 Quadratic Discriminant Analysis (QDA)

The quadratic discriminant classifier is also a statistical classification and machine learning technique, which separates two or more classes of objects by means of a quadratic surface. It is a more general version of the linear discriminant analysis classifier.

The QDA is closely related to LDA except for the fact that here measurements from each class are normally distributed. In QDA unlike in LDA, there is no assumption that the covariance of each class is identical. The other steps involved in QDA classification are the same as in the LDA.

4. Results

4.1 Detection of *Aspergillus flavus*

The results of the classification models developed for the detection of *A. flavus* infections in five types of pulses namely chick peas, green peas, lentils, pinto bean and kidney beans are presented in this section.

4.1.1 Chick peas

Figure 4.1, 4.2 and 4.3 show the mean reflectance, first factor PC loadings and second factor PC loadings, respectively, for healthy and *A. flavus* infected chick peas. Using first order PC loadings three significant wavelengths and using second order PC loadings two significant wavelengths were identified (Table 4.1). Using these significant wavelengths, statistical and histogram features were extracted and used in classification models. Classification accuracies (mean ± standard deviation, %) using six-way model for healthy and *A. flavus* infected chick peas using LDA and QDA are given in Table 4.2 and using two-way models are given in Table 4.3.

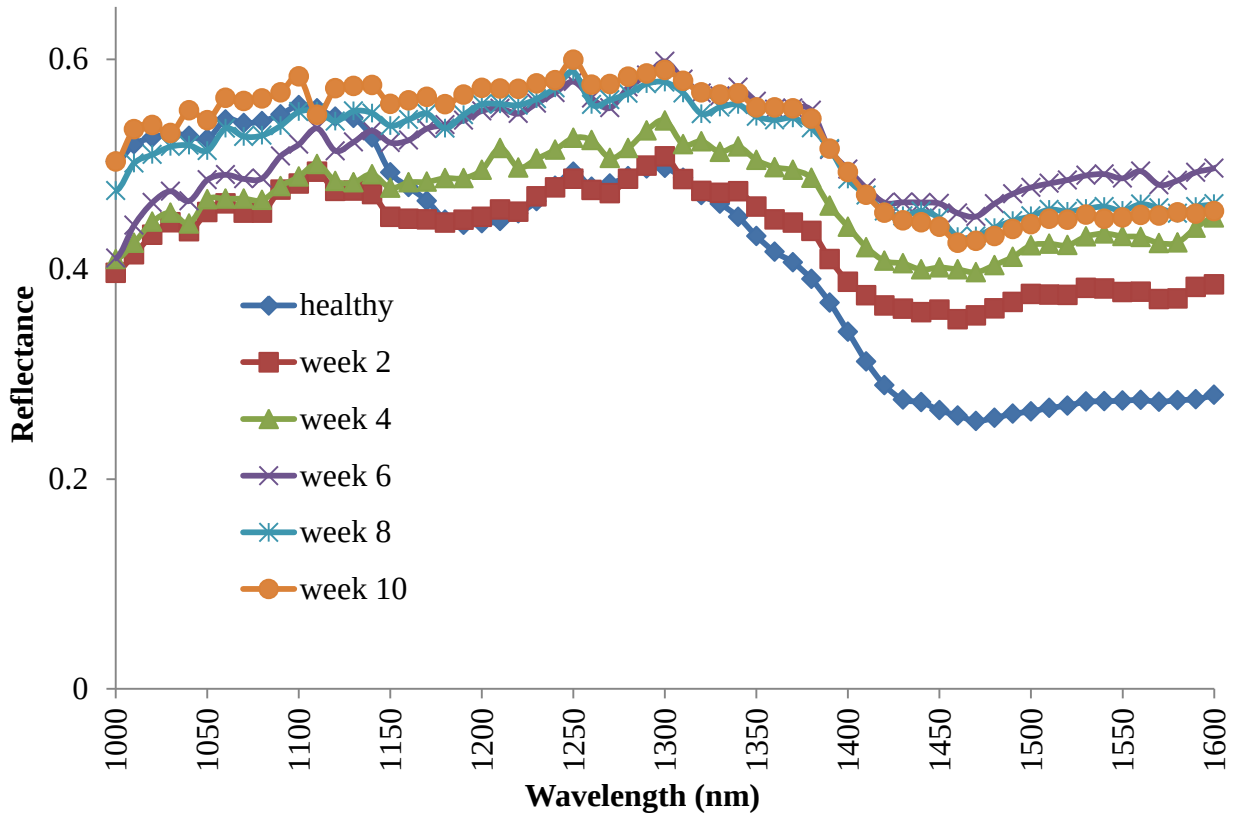

Fig. 4.1 Near-infrared mean reflectance spectra of healthy and *A. flavus* infected chick peas

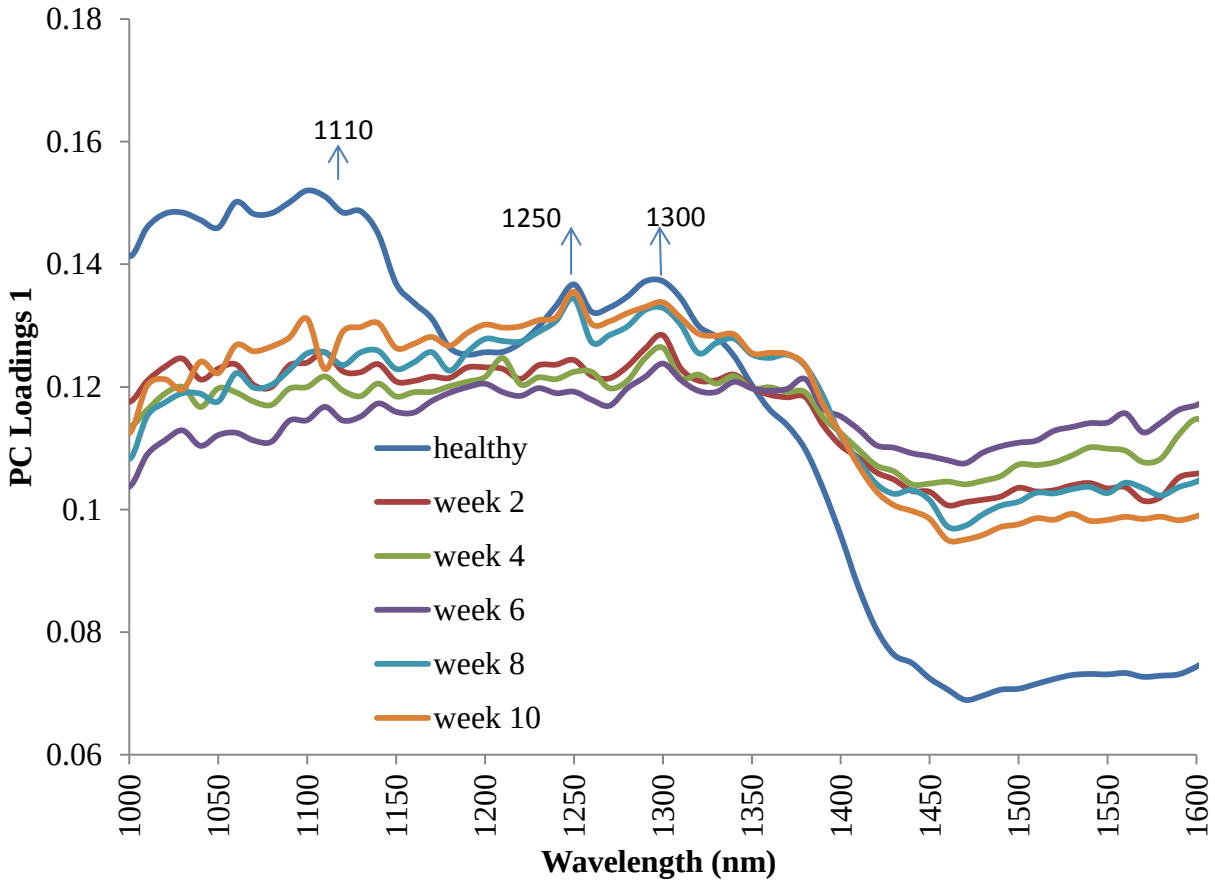

Fig. 4.2 First PC factor loadings of healthy and *A. flavus* infected chick peas for different weeks of infection stages

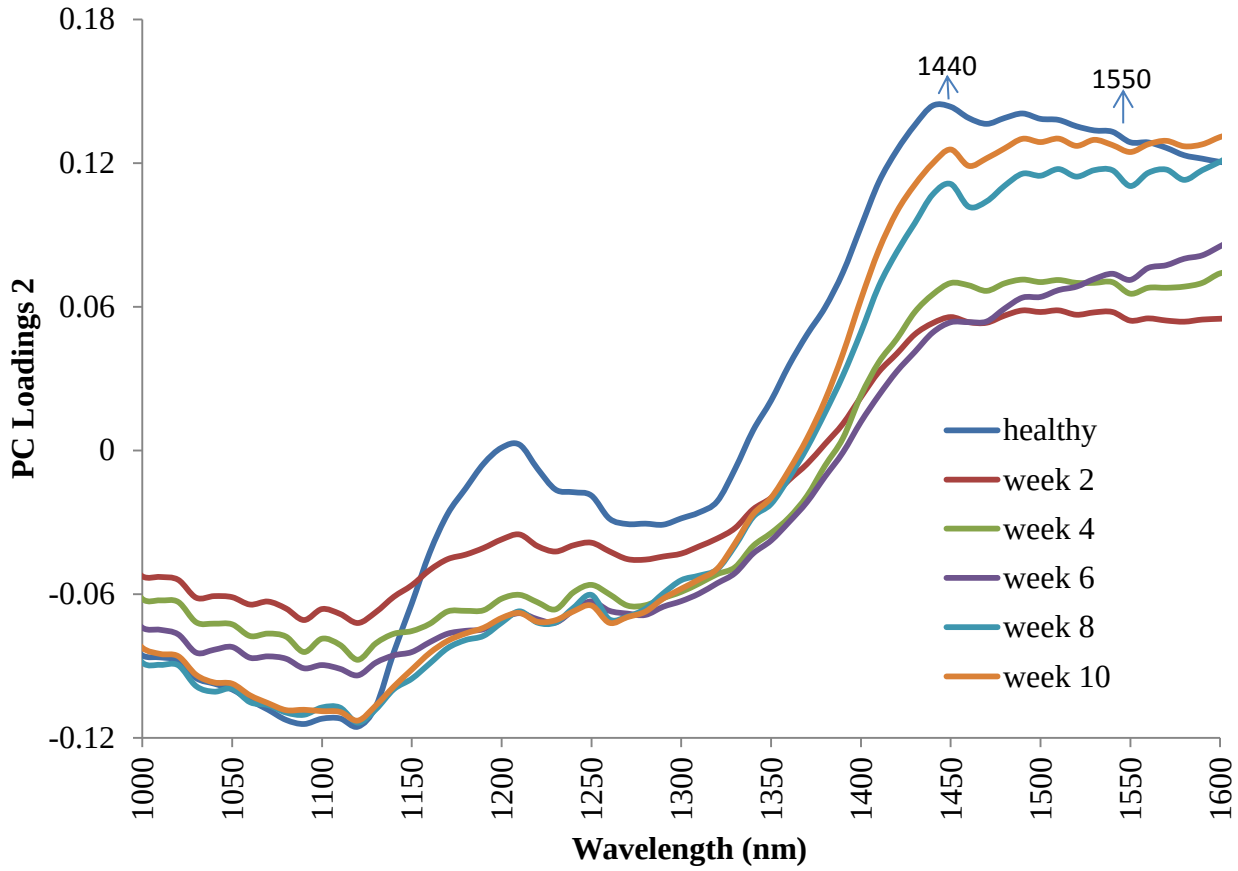

Fig. 4.3 Second PC factor loadings of healthy and *A. flavus* infected chick peas for different weeks of infection stages

Table 4.1 Identified wavelengths (nm) from the first and second principal component (PC) factor loadings for different pulse types

Pulse Type	First PC factor loading	Second PC factor loading
Chick peas	1110, 1250, 1300	1440, 1550
Green peas	1110, 1290, 1320	1440, 1550
Lentils	1110, 1250, 1300	1510, 1550
Pinto bean	1180, 1240, 1300	1420, 1520
Kidney bean	1020, 1110, 1290	1250, 1450

Table 4.2 Six-way classification accuracies (mean ± standard deviation, %) of *A. flavus* infected chick peas using LDA and QDA

	Healthy	Second week	Fourth week	Sixth week	Eighth week	Tenth week
LDA						
Healthy	**100.0±0.0**	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0
Second week	0.0±0.0	**96.2±2.1**	2.9±2.2	0.4±0.9	0.4±0.9	0.0±0.0
Fourth week	0.0±0.0	3.9±3.2	**85.2±4.7**	10.0±1.4	0.3±0.6	0.0±0.0
Sixth week	0.0±0.0	0.3±0.8	13.7±3.0	**84.4±2.9**	1.3±0.7	0.0±0.0
Eighth week	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	**95.5±2.0**	4.4±2.0
Tenth week	0.9±0.8	0.0±0.0	0.0±0.0	0.0±0.0	21.6±7.0	**77.4±6.4**
QDA						
Healthy	**100.0±0.0**	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0
Second week	0.2±0.6	**88.3±2.6**	8.4±3.8	1.2±1.2	0.7±0.9	0.9±0.8
Fourth week	0.0±0.0	4.3±4.4	**71.1±4.3**	10.0±1.4	0.6±0.9	0.0±0.0
Sixth week	0.0±0.0	0.2±0.6	10.3±2.7	**84.4±2.9**	0.9±1.4	0.0±0.0
Eighth week	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	**89.9±5.0**	10.0±5.2
Tenth week	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	11.5±4.4	**88.4±4.4**

Table 4.3 Two-way classification accuracies (mean ± standard deviation, %) of *A. flavus* infected chick peas using LDA and QDA

Sample	Statistical Classification using PC1 and PC2 Wavelengths	
	LDA	**QDA**
Healthy	100.0±0.0	100.0±0.0
A. flavus week 2	100.0±0.0	100.0±0.0
Healthy	100.0±0.0	100.0±0.0
A. flavus week 4	100.0±0.0	98.3±0.2
Healthy	100.0±0.0	100.0±0.0
A. flavus week 6	100.0±0.0	100.0±0.0
Healthy	100.0±0.0	100.0±0.0
A. flavus week 8	98.4±0.2	100.0±0.0
Healthy	100.0±0.0	100.0±0.0
A. flavus week 10	100.0±0.0	100.0±0.0

4.1.2 Green peas

Figure 4.4, 4.5 and 4.6 show the mean reflectance, first factor PC loadings and second factor PC loadings, respectively, for healthy and *A. flavus* infected green peas. Using the features extracted from the five significant wavelengths (Table 4.1), the classification accuracies (mean ± standard deviation, %) using six-way model for healthy and *A. flavus* infected green peas using LDA and QDA are given in Table 4.4 and using two-way models are given in Table 4.5.

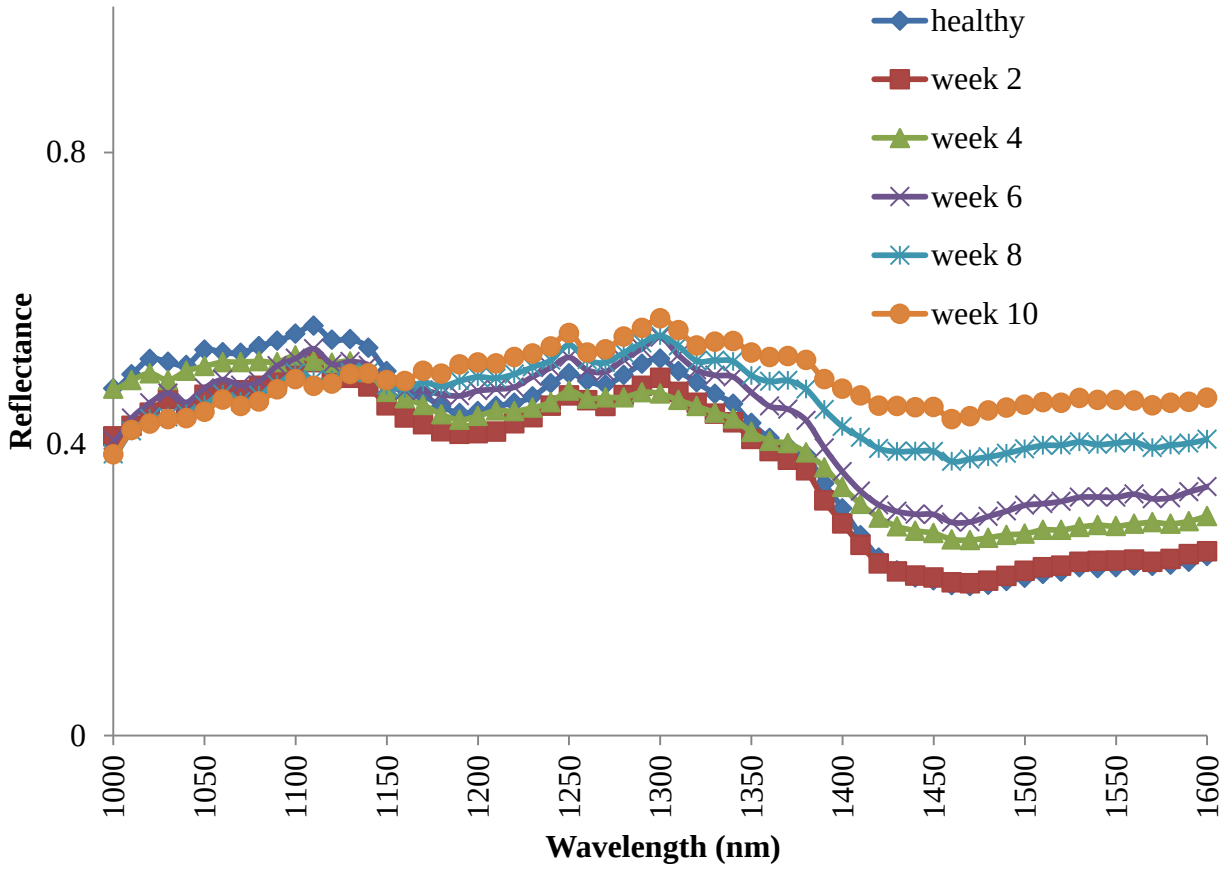

Fig. 4.4 Near-infrared mean reflectance spectra of healthy and *A. flavus* infected green peas

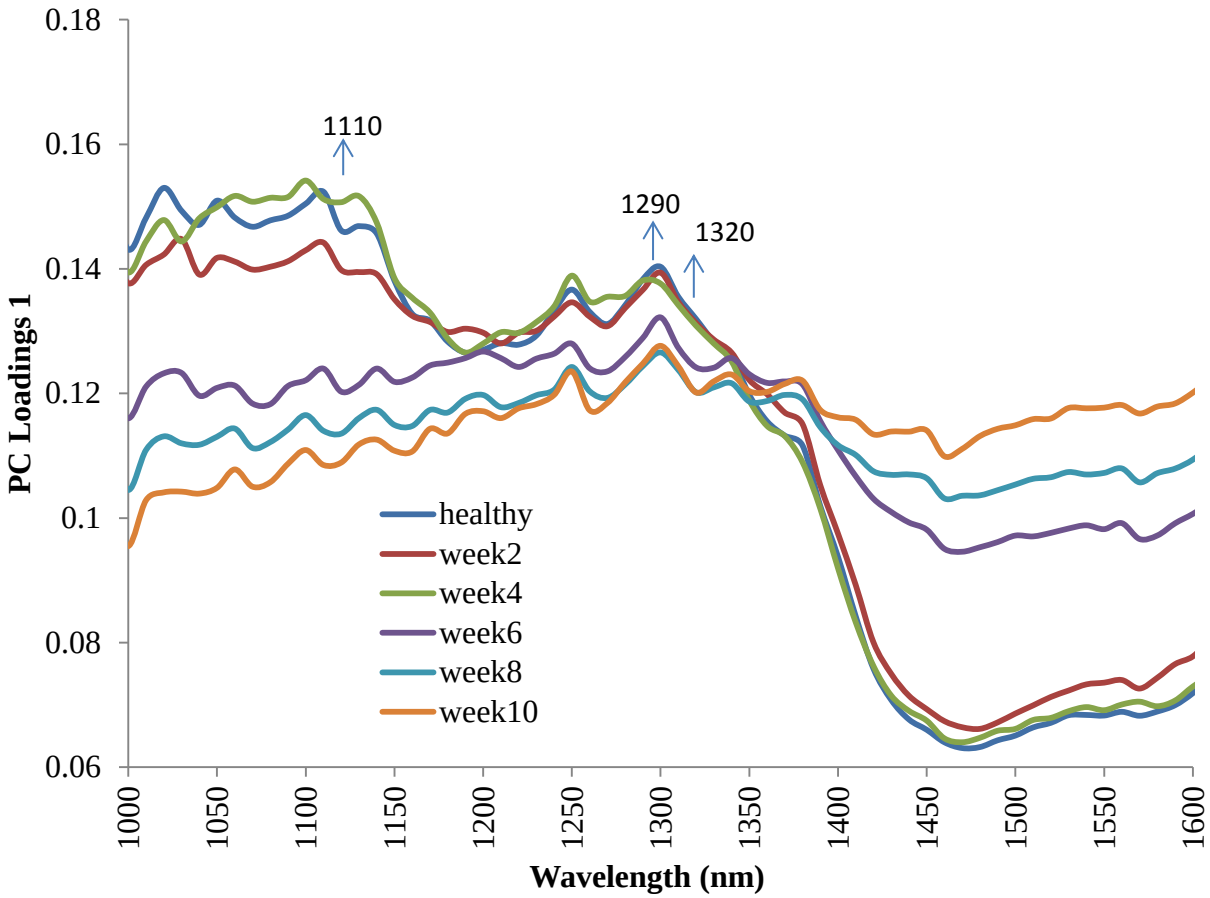

Fig. 4.5 First PC factor loadings of healthy and *A. flavus* infected green peas for different weeks of infection stages

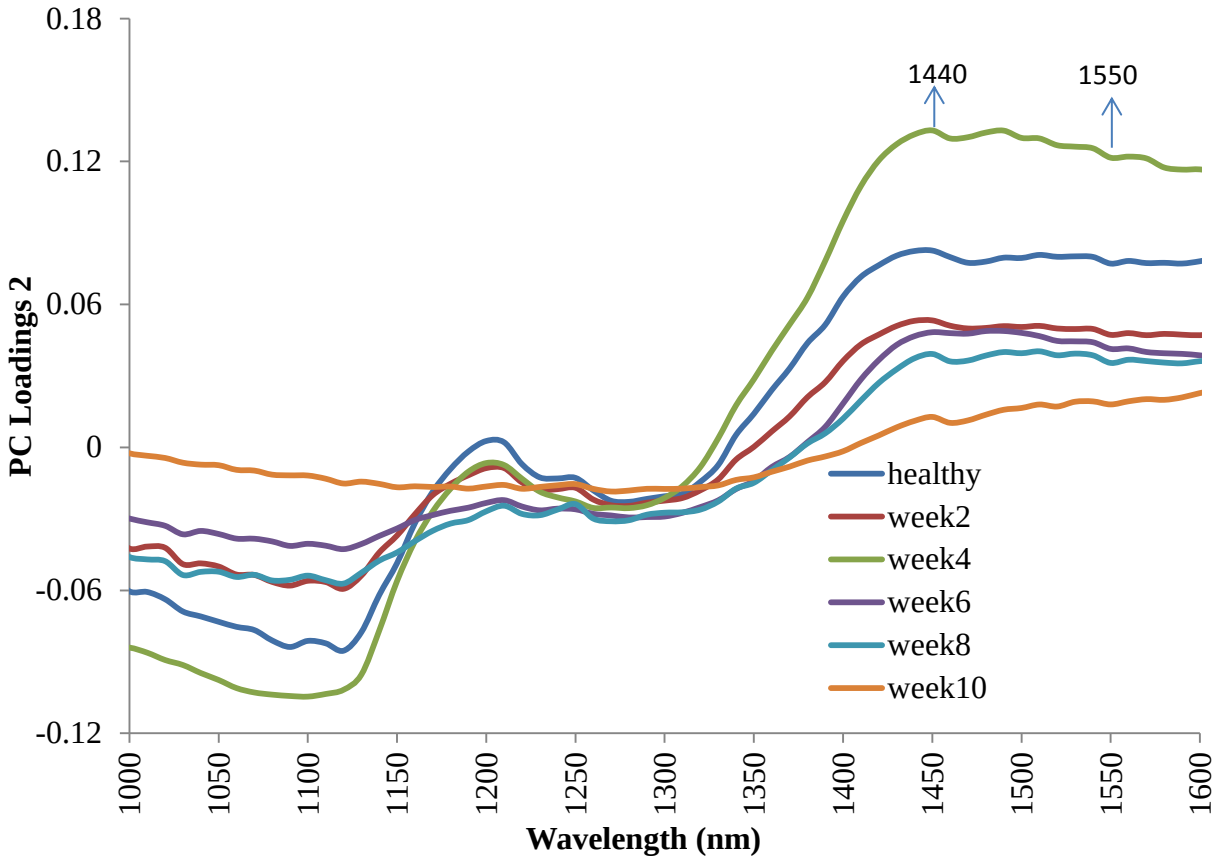

Fig. 4.6 Second PC factor loadings of healthy and *A. flavus* infected green peas for different weeks of infection stages

Table 4.4 Six-way classification accuracies (mean ± standard deviation, %) of *A. flavus* infected green peas using LDA and QDA

	Healthy	Second week	Fourth week	Sixth week	Eighth week	Tenth Week
LDA						
Healthy	**100.0±0.0**	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0
Second week	0.0±0.0	**98.2±1.6**	0.0±0.0	1.4±1.6	0.0±0.0	0.0±0.0
Fourth week	0.0±0.0	0.0±0.0	**99.6±0.7**	0.0±0.0	0.0±0.0	0.3±0.7
Sixth week	0.0±0.0	0.0±0.0	0.0±0.0	**98.3±1.1**	1.6±1.1	0.0±0.0
Eighth week	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	**99.7±0.6**	0.2±0.6
Tenth week	0.9±0.8	0.0±0.0	0.6±0.9	0.0±0.0	0.0±0.0	**99.3±0.9**
QDA						
Healthy	**98.1±2.2**	0.0±0.0	0.0±0.0	0.0±0.0	1.1±2.4	0.7±0.9
Second week	0.0±0.0	**99.6±0.6**	0.0±0.0	0.3±0.6	0.0±0.0	0.0±0.0
Fourth week	0.0±0.0	0.0±0.0	**99.3±0.9**	0.0±0.0	0.0±0.0	0.6±0.9
Sixth week	0.0±0.0	0.6±0.8	0.3±0.7	**99.0±0.8**	0.0±0.0	0.0±0.0
Eighth week	0.0±0.0	0.0±0.0	0.3±0.7	0.7±0.9	**98.9±0.9**	0.0±0.0
Tenth week	0.0±0.0	0.0±0.0	0.6±0.8	0.0±0.0	0.0±0.0	**99.3±0.8**

Table 4.5 Two-way classification accuracies (mean ± standard deviation, %) of *A. flavus* infected green peas using LDA and QDA

Sample	Statistical Classification using PC1 and PC2 Wavelengths	
	LDA	**QDA**
Healthy	100.0±0.0	100.0±0.0
A. flavus week 2	100.0±0.0	100.0±0.0
Healthy	100.0±0.0	100.0±0.0
A. flavus week 4	100.0±0.0	100.0±0.0
Healthy	100.0±0.0	100.0±0.0
A. flavus week 6	100.0±0.0	100.0±0.0
Healthy	100.0±0.0	97.3±0.2
A. flavus week 8	98.3±0.0	100.0±0.0
Healthy	100.0±0.0	100.0±0.0
A. flavus week 10	100.0±0.0	100.0±0.0

4.1.3 Lentils

Figure 4.7, 4.8 and 4.9 show the mean reflectance, first factor PC loadings and second factor PC loadings, respectively, for healthy and *A. flavus* infected lentils. Using the features extracted from the five significant wavelengths (Table 4.1), the classification accuracies (mean ± standard deviation, %) using six-way model for healthy and *A. flavus* infected lentils using LDA and QDA are given in Table 4.6 and using two-way models are given in Table 4.7.

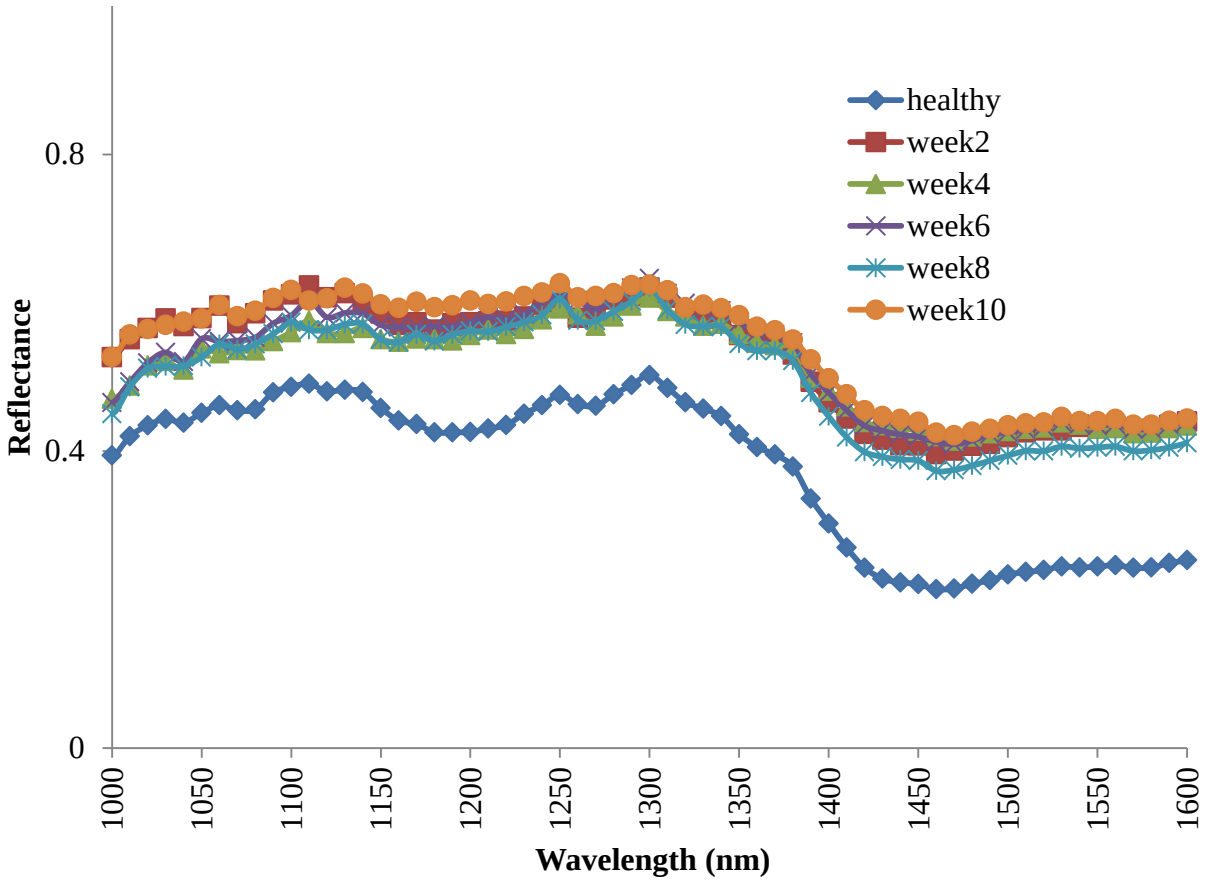

Fig. 4.7 Near-infrared mean reflectance spectra of healthy and infected (*A. flavus*) lentils

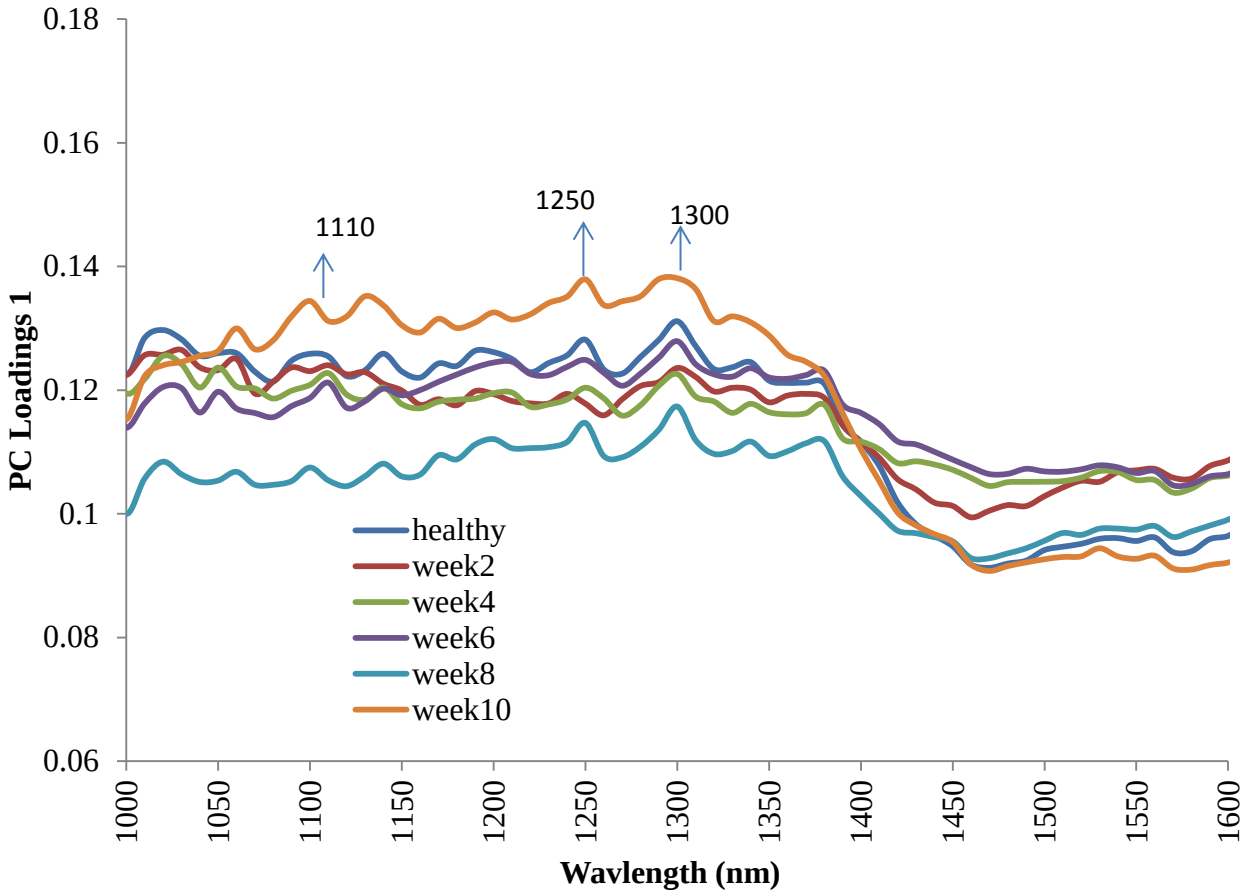

Fig. 4.8 First PC factor loadings of healthy and *A. flavus* infected lentils for different weeks of infection stages

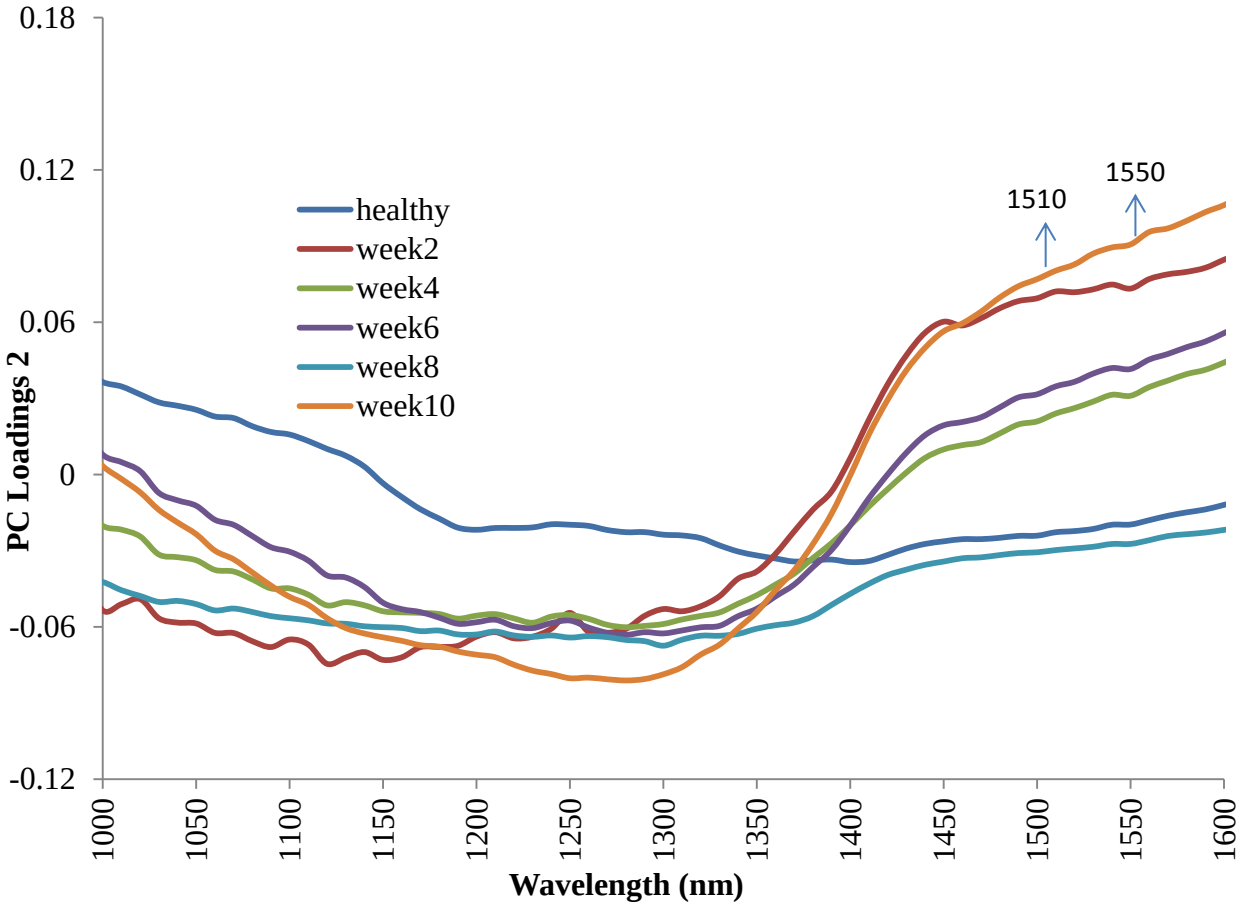

Fig. 4.9 Second PC factor loadings of healthy and *A. flavus* infected lentils for different weeks of infection stages

Table 4.6 Six-way classification accuracies (mean ± standard deviation, %) of *A. flavus* infected lentils using LDA and QDA

	Healthy	Second week	Fourth week	Sixth week	Eighth week	Tenth week
LDA						
Healthy	**99.3±0.9**	0.0±0.0	0.0±0.0	0.3±0.7	0.0±0.0	0.3±0.7
Second week	0.0±0.0	**99.7±0.5**	0.2±0.5	0.0±0.0	0.0±0.0	0.0±0.0
Fourth week	0.0±0.0	0.0±0.0	**77.8±6.3**	20.9±6.3	0.0±0.0	1.2±1.2
Sixth week	0.0±0.0	0.0±0.0	23.3±5.7	**71.6±5.3**	1.6±1.1	4.9±3.0
Eighth week	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	**100.0±0.0**	0.0±0.0
Tenth week	0.3±0.7	0.0±0.0	4.9±2.3	4.4±2.2	0.0±0.0	**90.2±3.1**
QDA						
Healthy	**99.3±0.9**	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.6±0.9
Second week	0.0±0.0	**100.0±0.0**	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0
Fourth week	0.9±2.1	0.0±0.0	**69.1±18.6**	22.9±13.1	0.0±0.0	6.9±4.6
Sixth week	0.0±0.0	0.0±0.0	11.1±9.6	**78.4±16.5**	0.0±0.0	10.3±7.6
Eighth week	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	**100.0±0.0**	0.0±0.0
Tenth week	0.0±0.0	0.0±0.0	3.8±3.0	2.5±3.3	0.0±0.0	**93.5±5.9**

Table 4.7 Two-way classification accuracies (mean ± standard deviation, %) of *A. flavus* infected lentils using LDA and QDA

Sample	Statistical Classification using PC1 and PC2 Wavelengths	
	LDA	**QDA**
Healthy	100.0±0.0	100.0±0.0
A. flavus week 2	100.0±0.0	100.0±0.0
Healthy	100.0±0.0	100.0±0.0
A. flavus week 4	100.0±0.0	98.4±0.6
Healthy	100.0±0.0	100.0±0.0
A. flavus week 6	98.4±0.3	100.0±0.0
Healthy	100.0±0.0	100.0±0.0
A. flavus week 8	100.0±0.0	100.0±0.0
Healthy	100.0±0.0	100.0±0.0
A. flavus week 10	100.0±0.0	100.0±0.0

4.1.4 Pinto bean

Figure 4.10, 4.11 and 4.12 show the mean reflectance, first factor PC loadings and second factor PC loadings, respectively, for healthy and *A. flavus* infected pinto beans. Using the features extracted from the five significant wavelengths (Table 4.1), the classification accuracies (mean ± standard deviation, %) using six-way model for healthy and *A. flavus* infected pinto beans using LDA and QDA are given in Table 4.8 and using two-way models are given in Table 4.9.

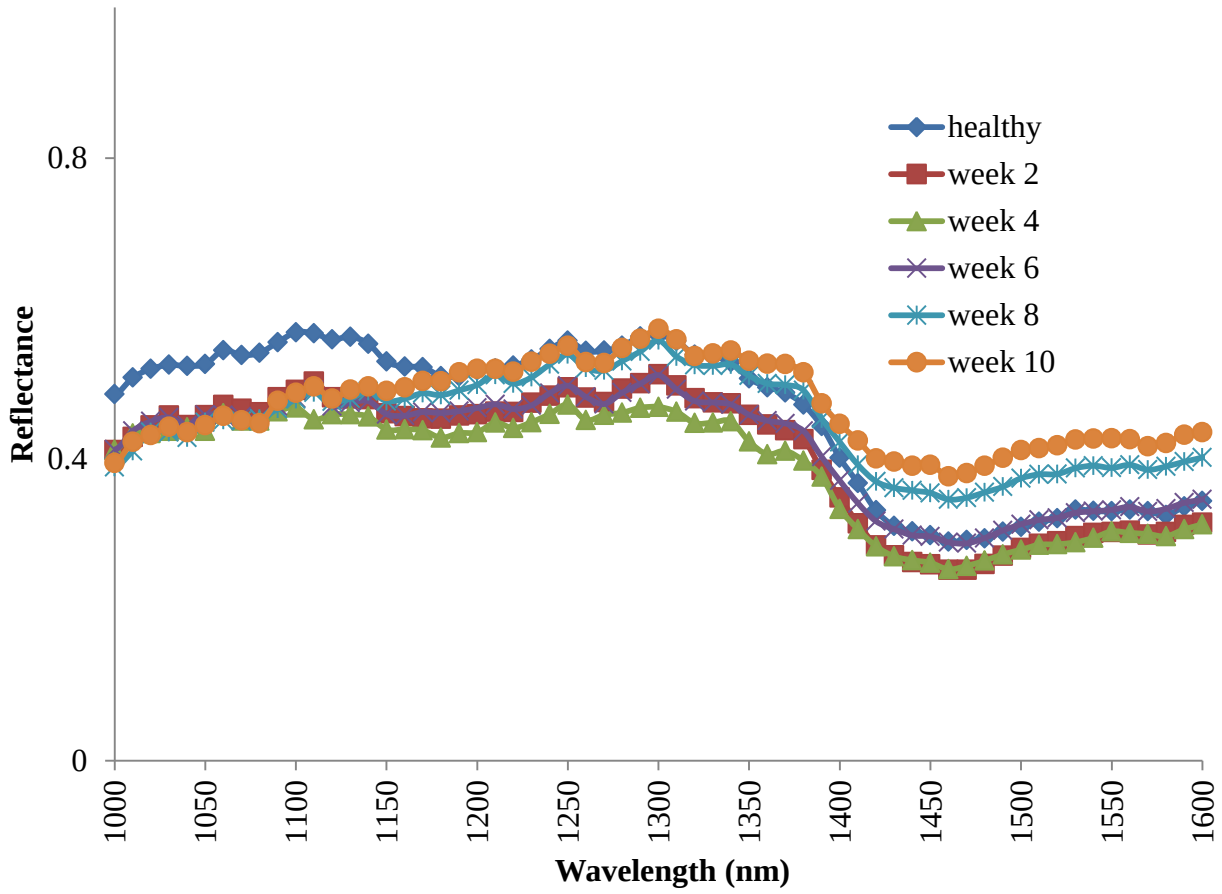

Fig. 4.10 Near-infrared mean reflectance spectra of healthy and (*A. flavus*) infected pinto bean

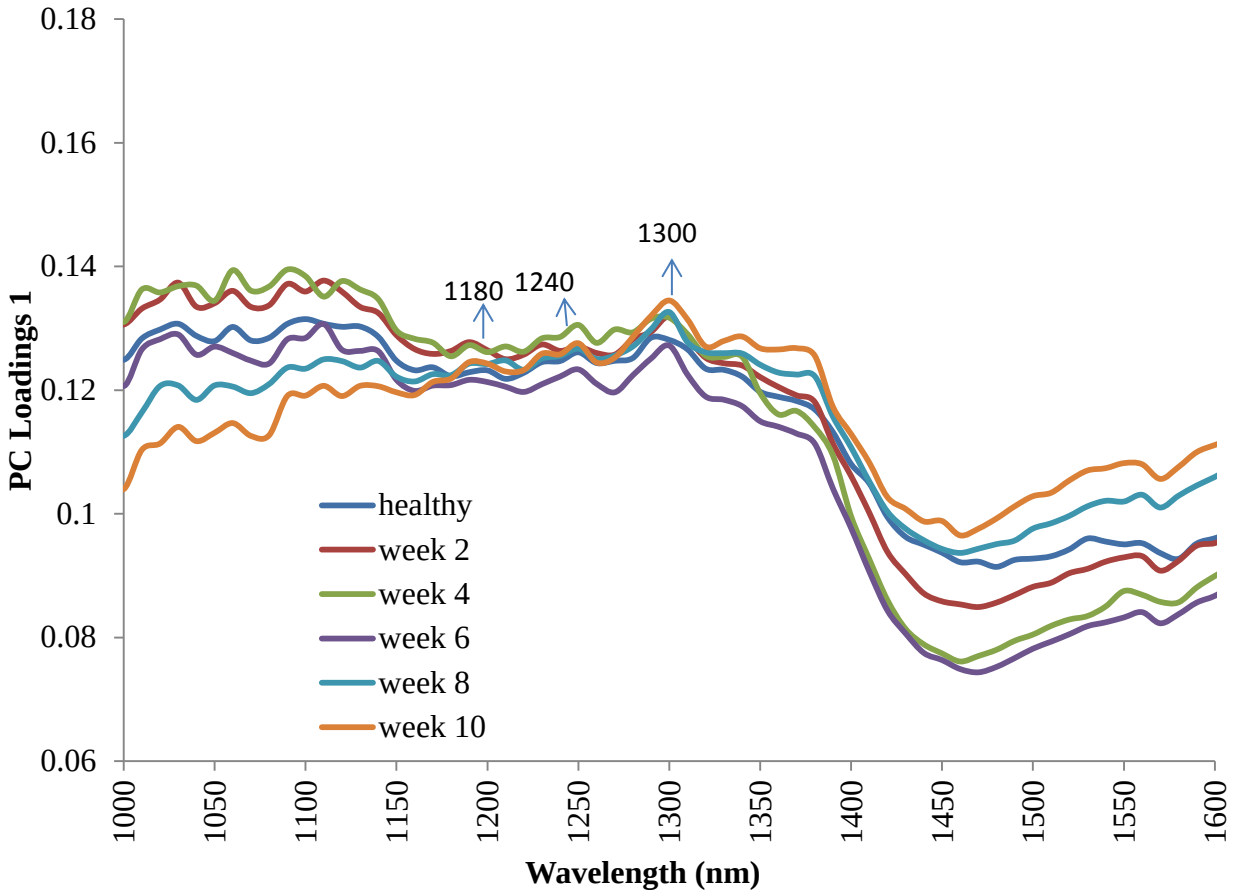

Fig. 4.11 First PC factor loadings of healthy and *A. flavus* infected pinto bean for different weeks of infection stages

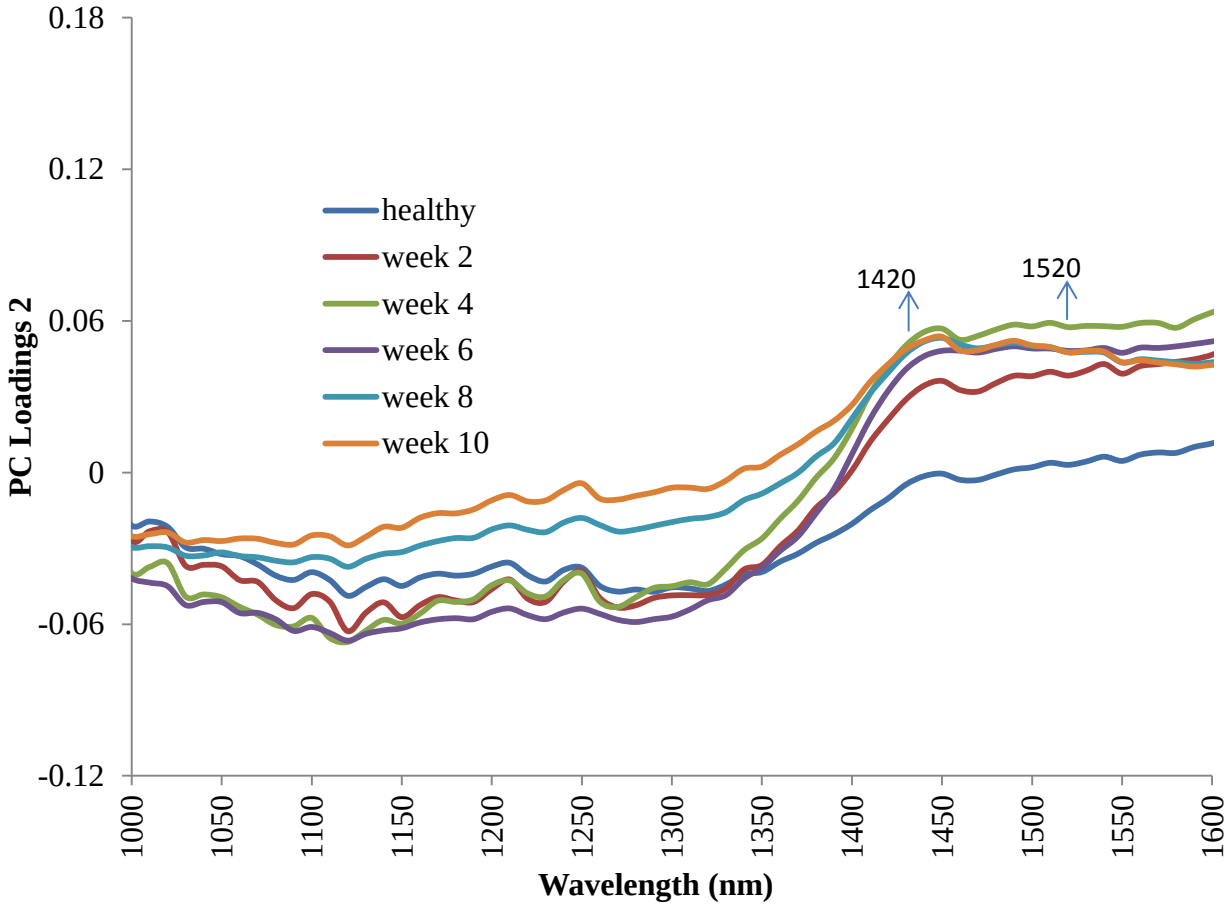

Fig. 4.12 Second PC factor loadings of healthy and *A. flavus* infected pinto bean for different weeks of infection stages

Table 4.8 Six-way classification accuracies (mean ± standard deviation, %) of *A. flavus* infected pinto bean using LDA and QDA

	Healthy	Second week	Fourth week	Sixth week	Eighth week	Tenth week
LDA						
Healthy	**100.0±0.0**	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0
Second week	0.0±0.0	**97.5±2.1**	2.1±2.4	0.3±0.7	0.0±0.0	0.0±0.0
Fourth week	0.0±0.0	2.4±2.8	**90.9±9.0**	6.3±5.7	0.3±0.6	0.0±0.0
Sixth week	0.0±0.0	0.3±0.8	6.9±6.4	**91.2±6.4**	1.4±1.8	0.0±0.0
Eighth week	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	**97.7±2.1**	0.0±0.0
Tenth week	0.6±0.8	0.0±0.0	0.3±0.7	0.0±0.0	11.0±10.5	**88.0±10.5**
QDA						
Healthy	**99.3±0.9**	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.6±0.9
Second week	0.0±0.0	**92.8±5.9**	5.1±4.8	0.9±0.8	0.3±0.8	0.6±0.8
Fourth week	0.9±2.1	3.9±4.7	**75.1±13.6**	18.9±10.6	0.3±0.7	1.6±2.8
Sixth week	0.0±0.0	0.2±0.6	7.7±4.7	**89.7±5.8**	0.2±0.6	1.9±4.2
Eighth week	0.0±0.0	0.0±0.0	0.0±0.0	0.3±0.8	**94.5±5.8**	5.0±6.2
Tenth week	0.0±0.0	0.0±0.0	1.0±2.3	0.6±1.5	8.3±8.1	**89.8±6.7**

Table 4.9 Two-way classification accuracies (mean ± standard deviation, %) of *A. flavus* infected pinto bean using LDA and QDA

Sample	Statistical Classification using PC1 and PC2 Wavelengths	
	LDA	**QDA**
Healthy	100.0±0.0	100.0±0.0
A. flavus week 2	100.0±0.0	100.0±0.0
Healthy	100.0±0.0	100.0±0.0
A. flavus week 4	100.0±0.0	98.4±3.4
Healthy	95.4±0.4	100.0±0.0
A. flavus week 6	98.4±1.5	100.0±0.0
Healthy	100.0±0.0	94.3±6.0
A. flavus week 8	100.0±0.0	92.8±0.0

4.1.5 Kidney bean

Figure 4.13, 4.14 and 4.15 show the mean reflectance, first factor PC loadings and second factor PC loadings, respectively, for healthy and *A. flavus* infected kidney beans. Using the features extracted from the five significant wavelengths (Table 4.1), the classification accuracies (mean ± standard deviation, %) using six-way model for healthy and *A. flavus* infected kidney beans using LDA and QDA are given in Table 4.10 and using two-way models are given in Table 4.11.

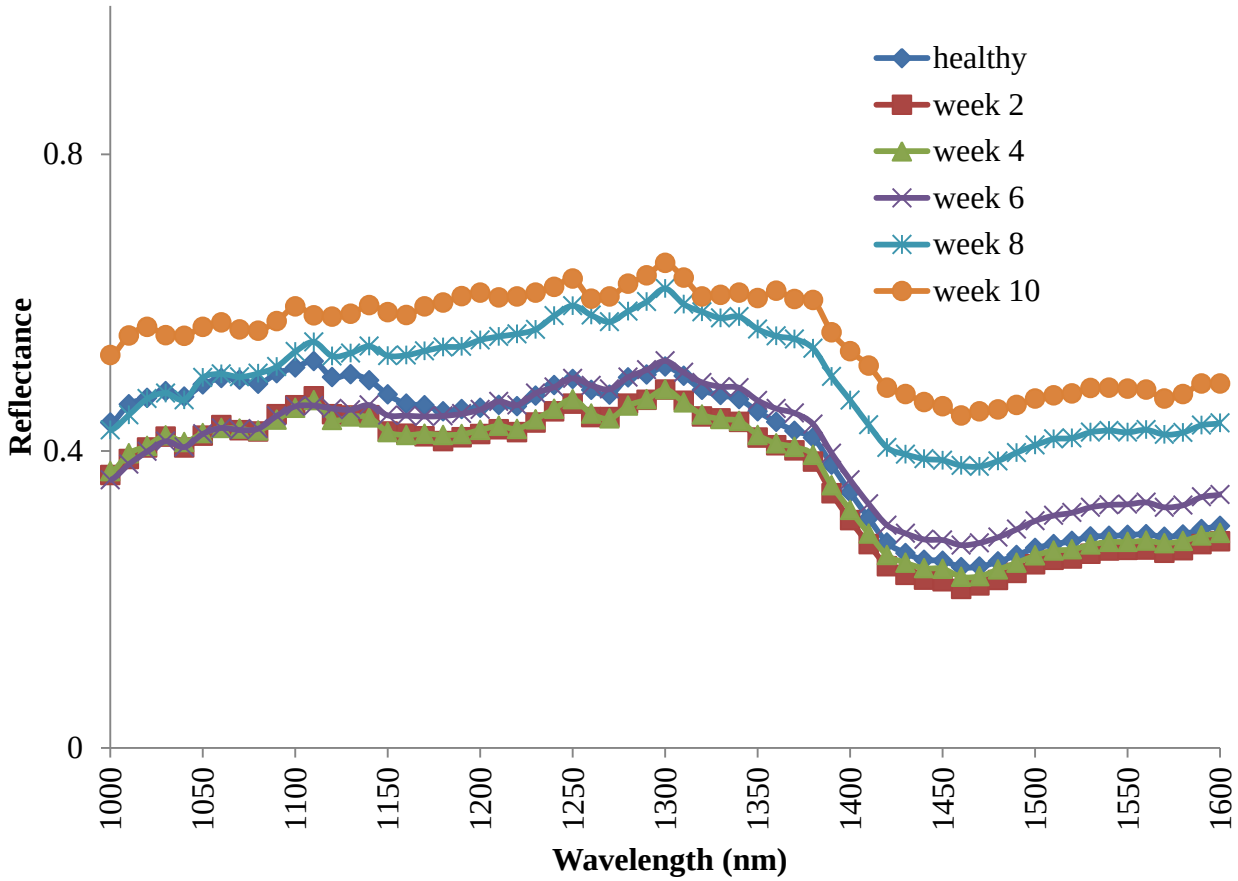

Fig. 4.13 Near-infrared mean reflectance spectra of healthy and infected *(A. flavus*) kidney bean

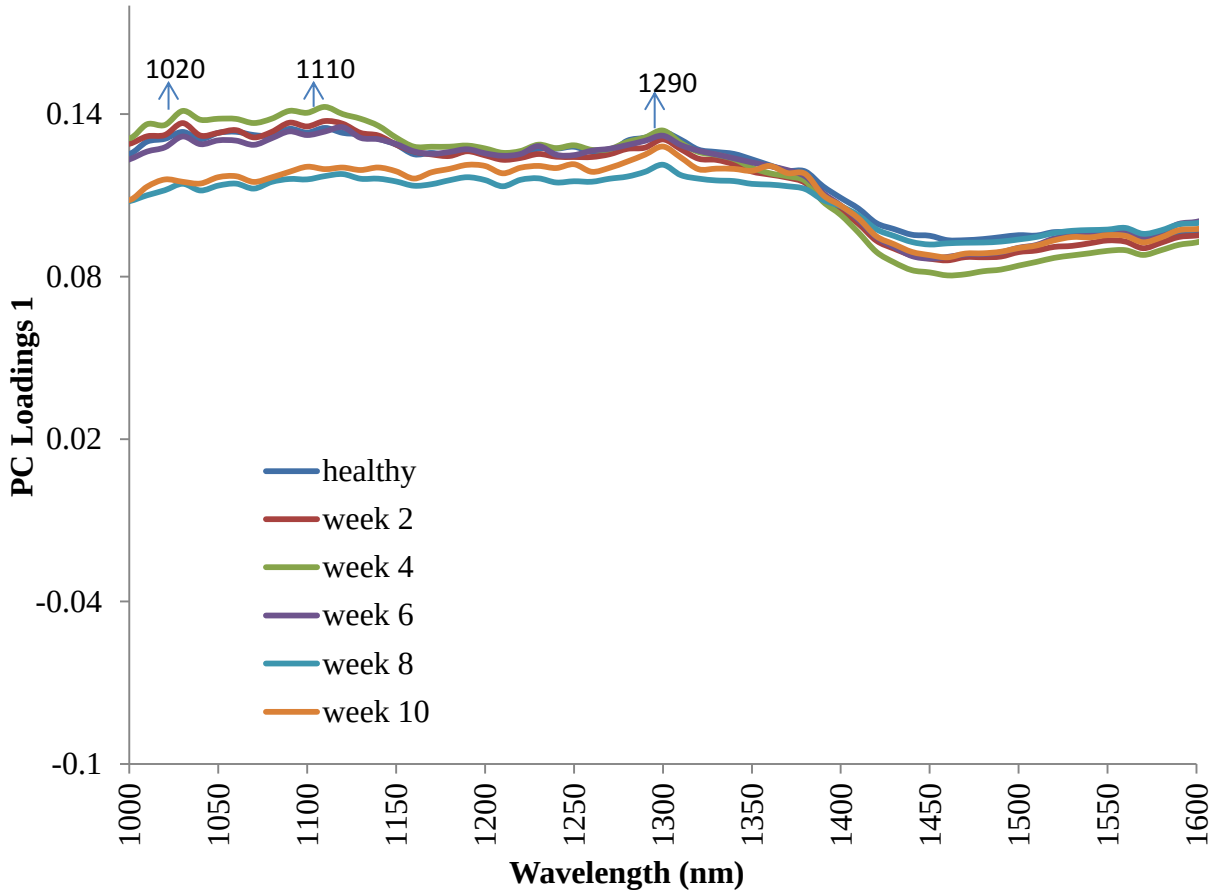

Fig. 4.14 First PC factor loadings of healthy and (*A. flavus*) infected kidney beans for different weeks of infection stages.

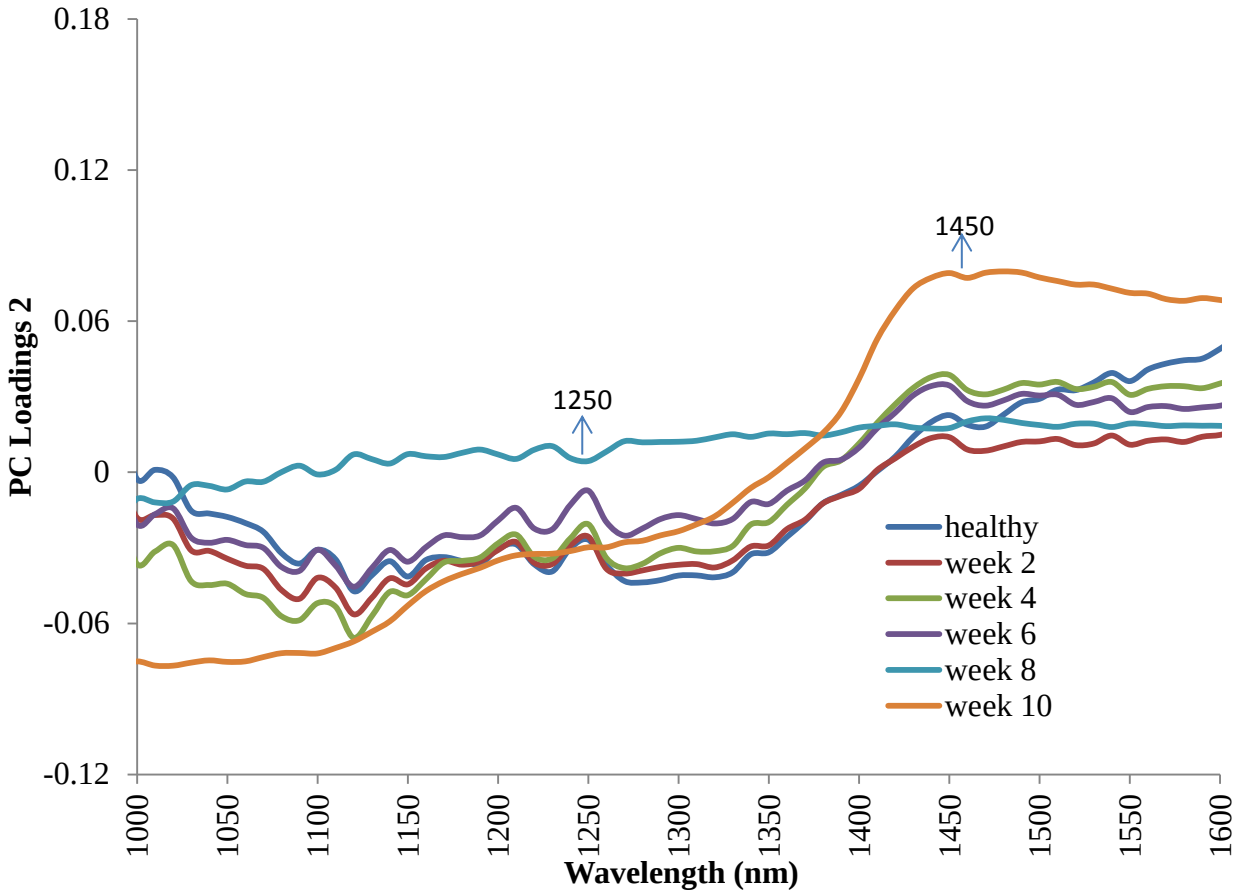

Fig. 4.15 Second PC factor loadings of healthy and (*A. flavus*) infected kidney beans for different weeks of infection stages.

Table 4.10 Six-way classification accuracies (mean ± standard deviation, %) of *A. flavus* infected kidney bean using LDA and QDA

	Healthy	Second week	Fourth week	Sixth week	Eighth week	Tenth week
LDA						
Healthy	**100.0±0.0**	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0
Second week	0.0±0.0	**98.5±2.1**	0.6±1.3	0.8±1.7	0.0±0.0	0.0±0.0
Fourth week	0.0±0.0	1.2±2.7	**95.5±8.9**	2.5±5.5	0.3±0.6	0.3±0.7
Sixth week	0.0±0.0	0.0±0.0	2.3±5.1	**96.3±5.5**	1.2±0.7	0.0±0.0
Eighth week	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	**99.0±1.4**	0.9±1.4
Tenth week	0.3±0.7	0.0±0.0	0.3±0.7	0.0±0.0	3.6±8.0	**95.7±8.6**
QDA						
Healthy	**98.5±2.4**	0.0±0.0	0.0±0.0	0.0±0.0	1.1±2.4	0.3±0.7
Second week	0.2±0.6	**97.8±3.8**	0.5±1.3	0.9±1.3	0.0±0.0	0.2±0.6
Fourth week	0.0±0.0	0.0±0.0	**87.1±16.9**	10.5±14.1	0.0±0.0	2.2±3.3
Sixth week	0.0±0.0	0.2±0.6	7.1±10.4	**87.9±18.4**	0.6±1.5	3.8±8.6
Eighth week	0.0±0.0	0.0±0.0	0.3±0.7	0.3±0.8	**96.1±6.5**	3.1±6.9
Tenth week	0.0±0.0	0.0±0.0	1.2±1.9	0.0±0.0	1.7±3.8	**97.0±3.6**

Table 4.11 Two-way classification accuracies (mean ± standard deviation, %) of *A. flavus* infected kidney bean using LDA and QDA

Sample	Statistical Classification using PC1 and PC2 Wavelengths	
	LDA	**QDA**
Healthy	100.0±0.0	100.0±0.0
A. flavus week 2	95.0±1.0	92.0±5.2
Healthy	100.0±0.0	100.0±0.0
A. flavus week 4	100.0±0.0	98.4±3.1
Healthy	100.0±0.0	100.0±0.0
A. flavus week 6	98.4±2.2	100.0±0.0
Healthy	98.0±0.7	100.0±0.0
A. flavus week 8	100.0±0.0	98.0±5.0
Healthy	100.0±0.0	100.0±0.0
A. flavus week 10	100.0±0.0	100.0±0.0

4.2 Detection of *Penicillium commune*

The results of the classification models developed for the detection of *Penicillium commune* infections in five types of pulses namely chick peas, green peas, lentils, pinto bean and kidney beans are presented in this section.

4.2.1 Chick peas

Figure 4.16, 4.17 and 4.18 show the mean reflectance, first factor PC loadings and second factor PC loadings, respectively, for healthy and *Penicillium commune* infected chick peas. Using the features extracted from the five significant wavelengths (Table 4.12), the classification accuracies (mean ± standard deviation, %) using six-way model for healthy and *Penicillium commune* infected chick peas using LDA and QDA are given in Table 4.13 and using two-way models are given in Table 4.14.

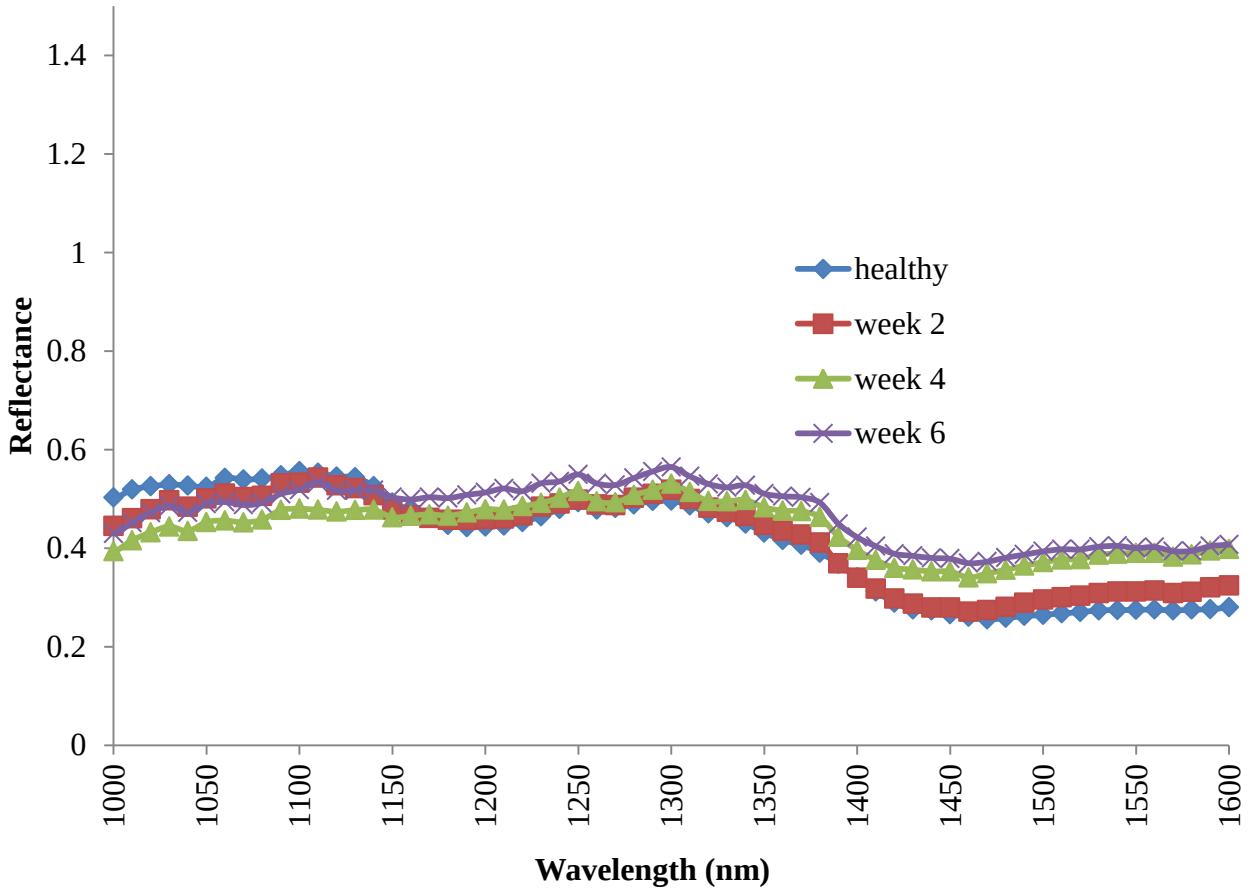

Fig. 4.16 Near-infrared mean reflectance spectra of healthy and infected (*Penicillium commune*) chick peas

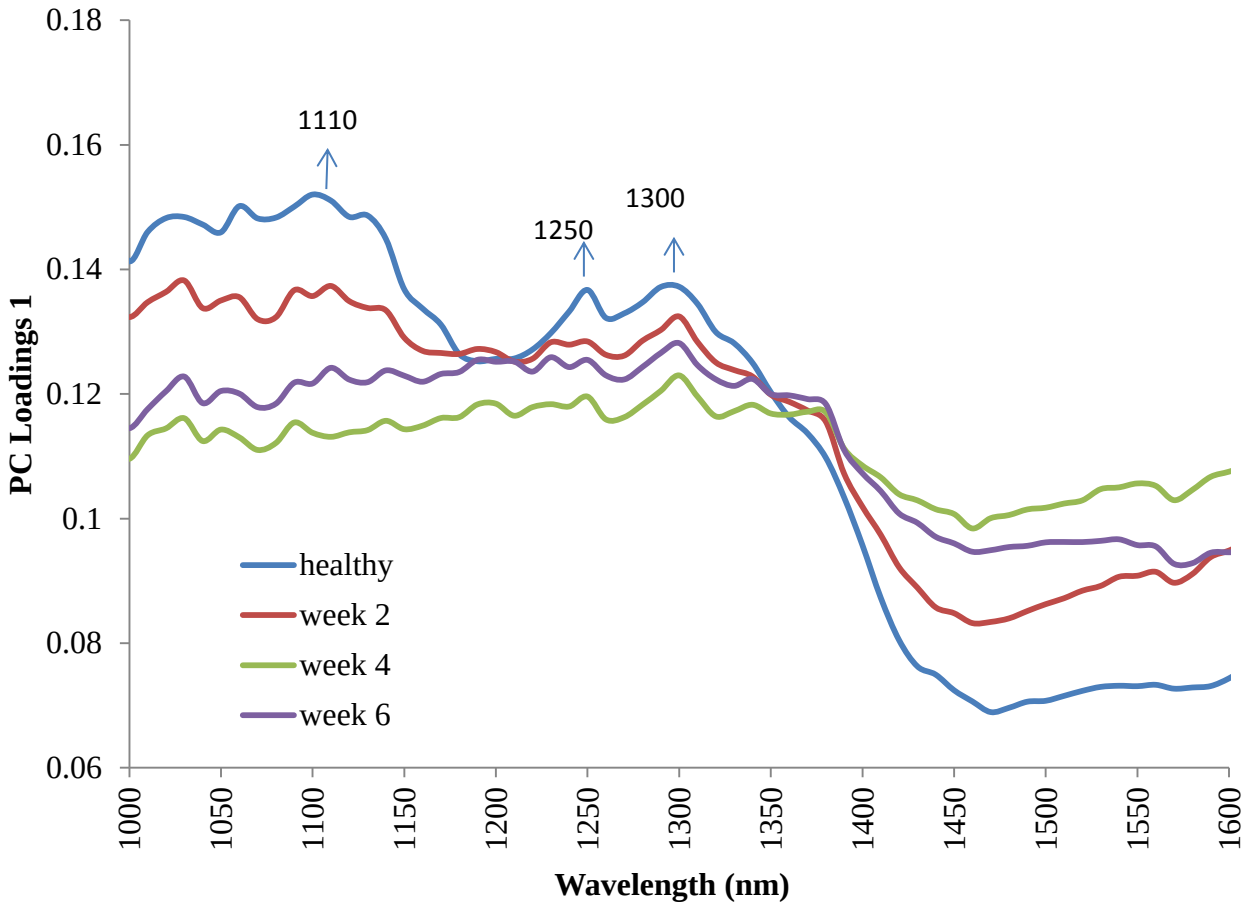

Fig. 4.17 First PC factor loadings of healthy and (*Penicillium commune*) infected chick peas for different weeks of infection stages

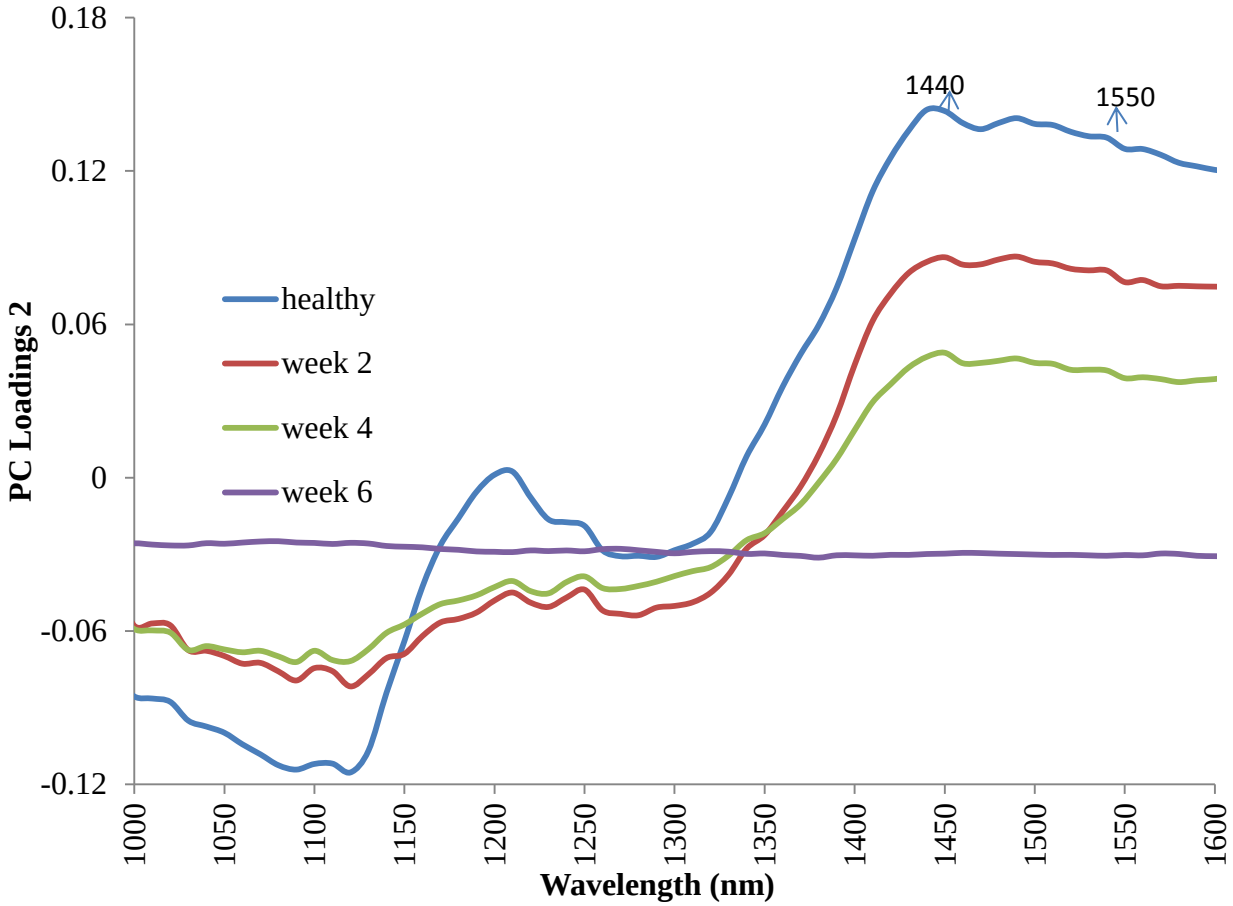

Fig. 4.18 Second PC factor loadings of healthy and (*Penicillium commune*) infected chick peas for different weeks of infection stages

Table 4.12 Identified wavelengths (nm) from the first and second principal component (PC) factor loadings for different pulse types

Pulse Type	First PC factor loading	Second PC factor loading
Chick peas	1110, 1250, 1300	1440, 1550
Green peas	1110, 1250, 1300	1440, 1550
Lentils	1220, 1250, 1300	1010, 1060
Pinto bean	1030, 1110, 1300	1250, 1450
Kidney bean	1020, 1110, 1200	1250, 1440

Table 4.13 Four-way classification accuracies (mean ± standard deviation, %) of *Penicillium commune* infected chick peas using LDA and QDA

	Healthy	Second week	Fourth week	Sixth week
LDA				
Healthy	100.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0
Second week	0.0±0.0	99.3±0.9	0.3±0.7	0.3±0.7
Fourth week	0.0±0.0	0.6±0.9	89.5±2.3	9.8±2.4
Sixth week	0.0±0.0	0.6±0.8	12.7±3.6	86.6±4.2
QDA				
Healthy	100.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0
Second week	0.6±1.4	95.8±3.4	0.0±0.0	3.5±3.1
Fourth week	0.0±0.0	0.3±0.7	94.8±2.5	4.8±2.1
Sixth week	0.0±0.0	0.0±0.0	14.0±6.8	85.9±6.8

Table 4.14 Two-way classification accuracies (mean ± standard deviation, %) of *Penicillium commune* infected chick peas using LDA and QDA

Sample	Statistical Classification using PC1 and PC2 Wavelengths	
	LDA	**QDA**
Healthy	100.0±0.0	100.0±0.0
P.commune week 2	100.0±0.0	100.0±0.0
Healthy	100.0±0.0	100.0±0.0
P.commune week 4	100.0±0.0	98.7±2.0
Healthy	100.0±0.0	100.0±0.0
P.commune week 6	100.0±0.0	100.0±0.0

4.2.2 Green peas

Figure 4.19, 4.20 and 4.21 show the mean reflectance, first factor PC loadings and second factor PC loadings, respectively, for healthy and *Penicillium commune* infected green peas. Using the features extracted from the five significant wavelengths (Table 4.12), the classification accuracies (mean ± standard deviation, %) using six-way model for healthy and *Penicillium commune* infected green peas using LDA and QDA are given in Table 4.15 and using two-way models are given in Table 4.16.

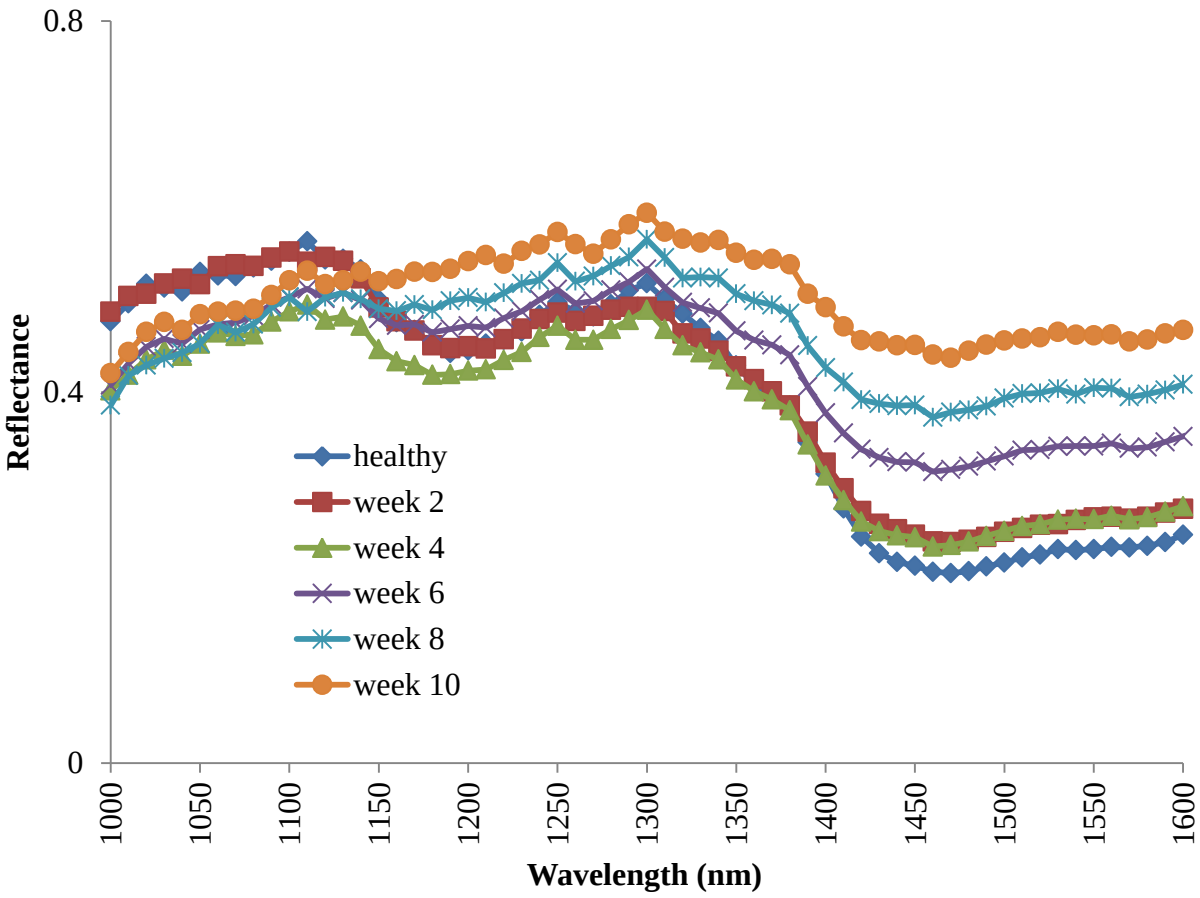

Fig. 4.19 Near-infrared mean reflectance spectra of healthy and infected (*Penicillium commune*) green peas

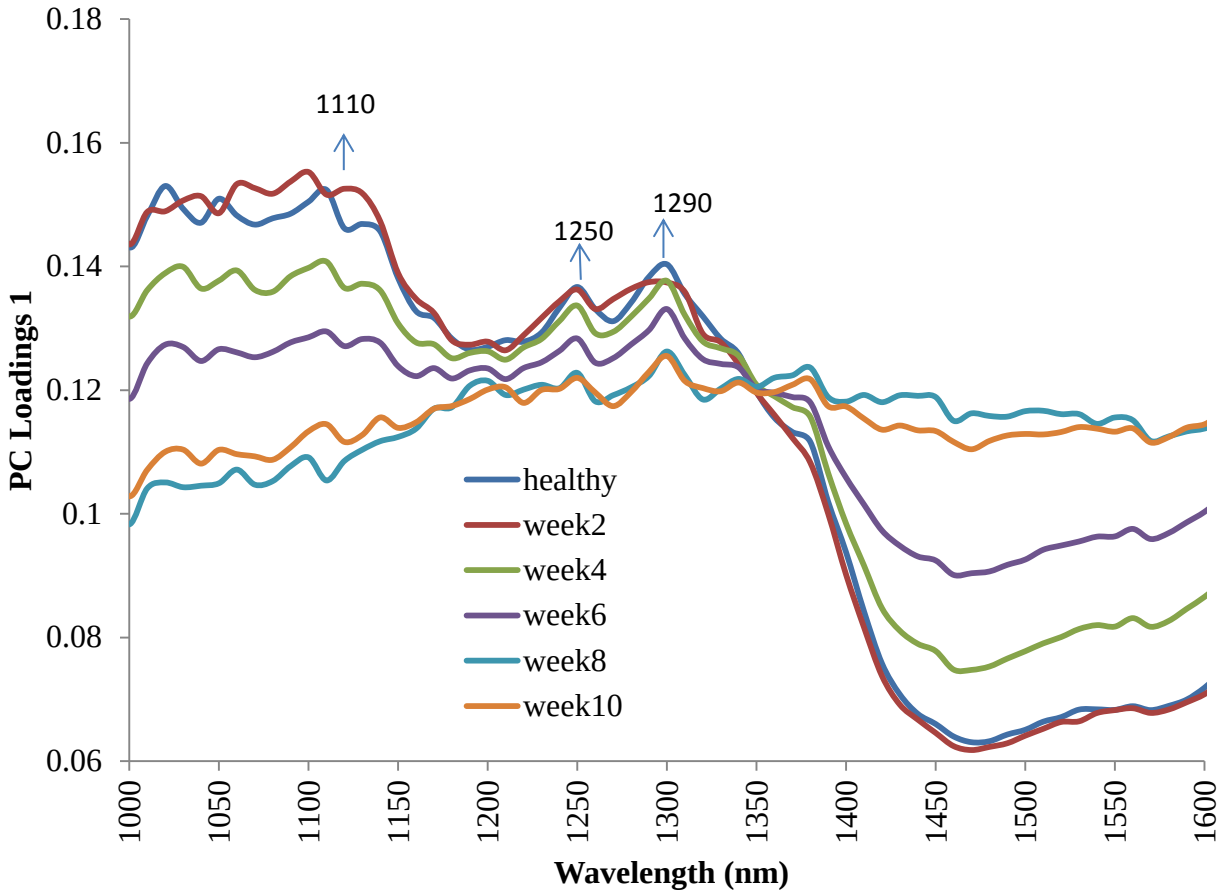

Fig. 4.20 First PC factor loadings of healthy and (*Penicillium commune*) infected green peas for different weeks of infection stages

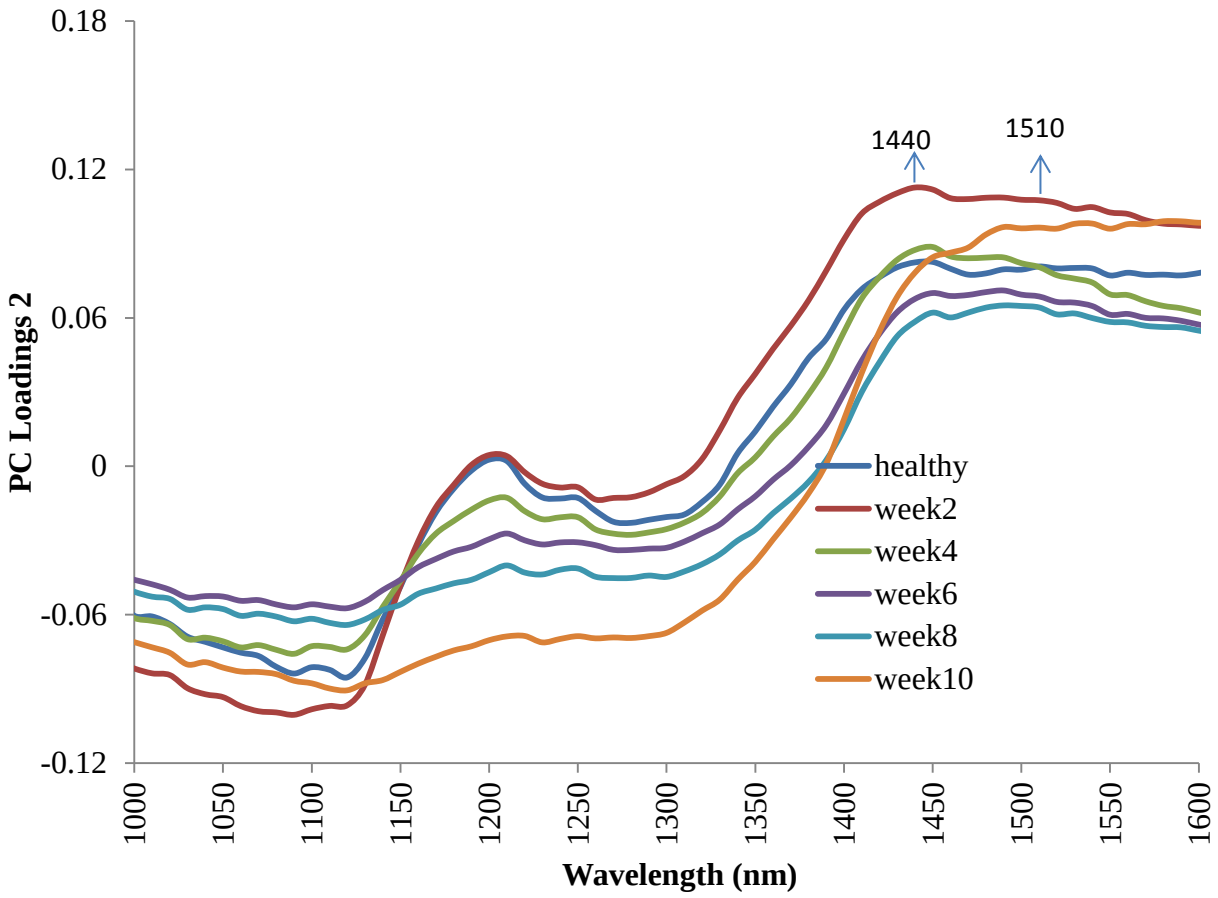

Fig. 4.21 Second PC factor loadings of healthy and (*Penicillium commune*) infected green peas for different weeks of infection stages

Table 4.15 Six-way classification accuracies (mean ± standard deviation, %) of *Penicillium commune* infected green peas using LDA and QDA

	Healthy	Second week	Fourth week	Sixth week	Eighth week	Tenth Week
LDA						
Healthy	**98.4±2.2**	0.0±0.0	1.5±2.2	0.0±0.0	0.0±0.0	0.0±0.0
Second week	0.0±0.0	**97.6±1.0**	1.9±1.4	0.3±0.7	0.0±0.0	0.0±0.0
Fourth week	0.9±2.1	6.9±2.0	**75.5±3.0**	15.7±4.8	0.6±1.5	0.0±0.0
Sixth week	0.0±0.0	0.0±0.0	11.1±5.5	**72.7±5.4**	15.0±5.7	1.0±0.9
Eighth week	0.0±0.0	0.0±0.0	0.6±0.9	18.2±7.4	**70.9±6.3**	10.0±4.3
Tenth week	0.0±0.0	0.0±0.0	0.0±0.0	6.6±1.6	12.5±8.0	**80.7±7.3**
QDA						
Healthy	**97.6±2.5**	0.2±0.6	2.0±2.3	0.0±0.0	0.0±0.0	0.0±0.0
Second week	0.0±0.0	**94.5±6.0**	4.0±4.3	0.6±1.9	0.6±0.9	0.0±0.0
Fourth week	0.0±0.0	4.6±3.7	**66.1±4.9**	27.4±6.7	0.9±1.4	0.6±1.5
Sixth week	0.0±0.0	0.2±0.6	10.9±2.9	**71.1±2.5**	15.0±4.0	3.1±1.4
Eighth week	0.0±0.0	0.4±0.9	1.1±1.0	9.8±2.9	**75.7±3.9**	12.9±4.4
Tenth week	0.0±0.0	0.0±0.0	0.6±0.8	1.2±1.2	11.8±3.0	**86.2±2.1**

Table 4.16 Two-way classification accuracies (mean ± standard deviation, %) of *Penicillium commune* infected green peas using LDA and QDA

Sample	Statistical Classification using PC1 and PC2 Wavelengths	
	LDA	**QDA**
Healthy	100.0±0.0	100.0±0.0
P.commune week 2	100.0±0.0	100.0±0.0
Healthy	100.0±0.0	100.0±0.0
P.commune week 4	100.0±0.0	100.0±0.0
Healthy	100.0±0.0	100.0±0.0
P.commune week 6	100.0±0.0	100.0±0.0
Healthy	100.0±0.0	100.0±0.0
P.commune week 8	100.0±0.0	98.2±2.0
Healthy	100.0±0.0	100.0±0.0
P.commune week 10	100.0±0.0	100.0±0.0

4.2.3 Lentils

Figure 4.22, 4.23 and 4.24 show the mean reflectance, first factor PC loadings and second factor PC loadings, respectively, for healthy and *Penicillium commune* infected lentils. Using the features extracted from the five significant wavelengths (Table 4.12), the classification accuracies (mean ± standard deviation, %) using six-way model for healthy and *Penicillium commune* infected lentils using LDA and QDA are given in Table 4.17 and using two-way models are given in Table 4.18.

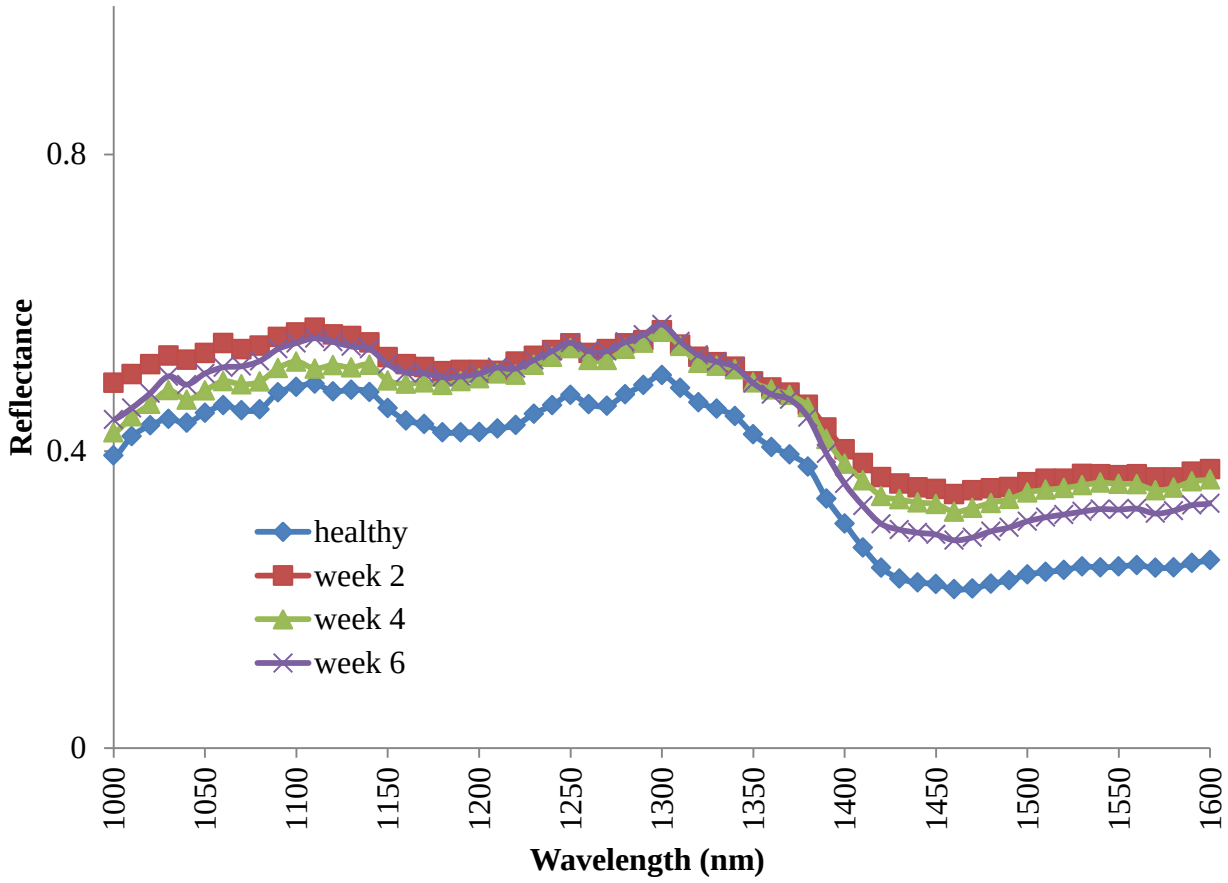

Fig. 4.22 Near-infrared mean reflectance spectra of healthy and infected (*Penicillium commune*) lentils.

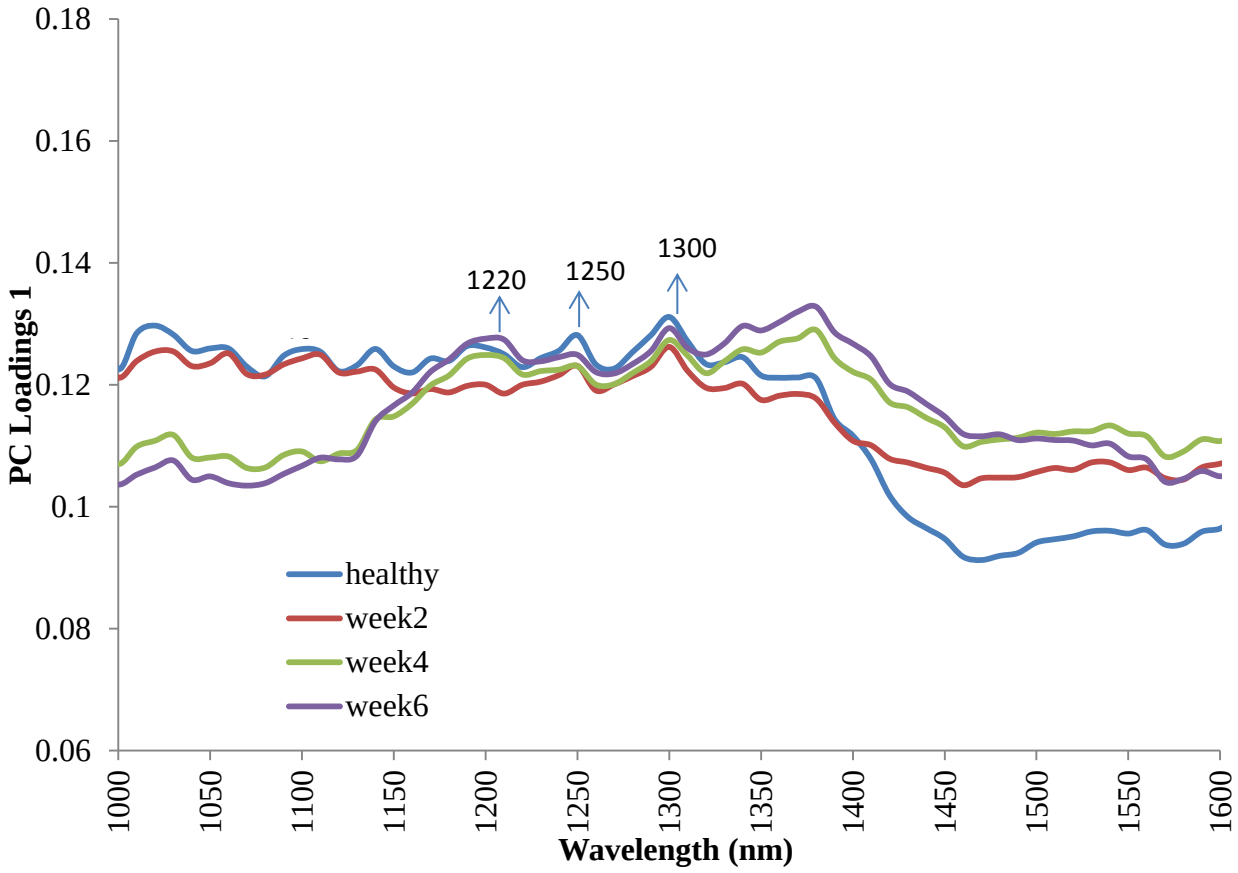

Fig. 4.23 First PC factor loadings of healthy and (*Penicillium commune*) infected lentils for different weeks of infection stages

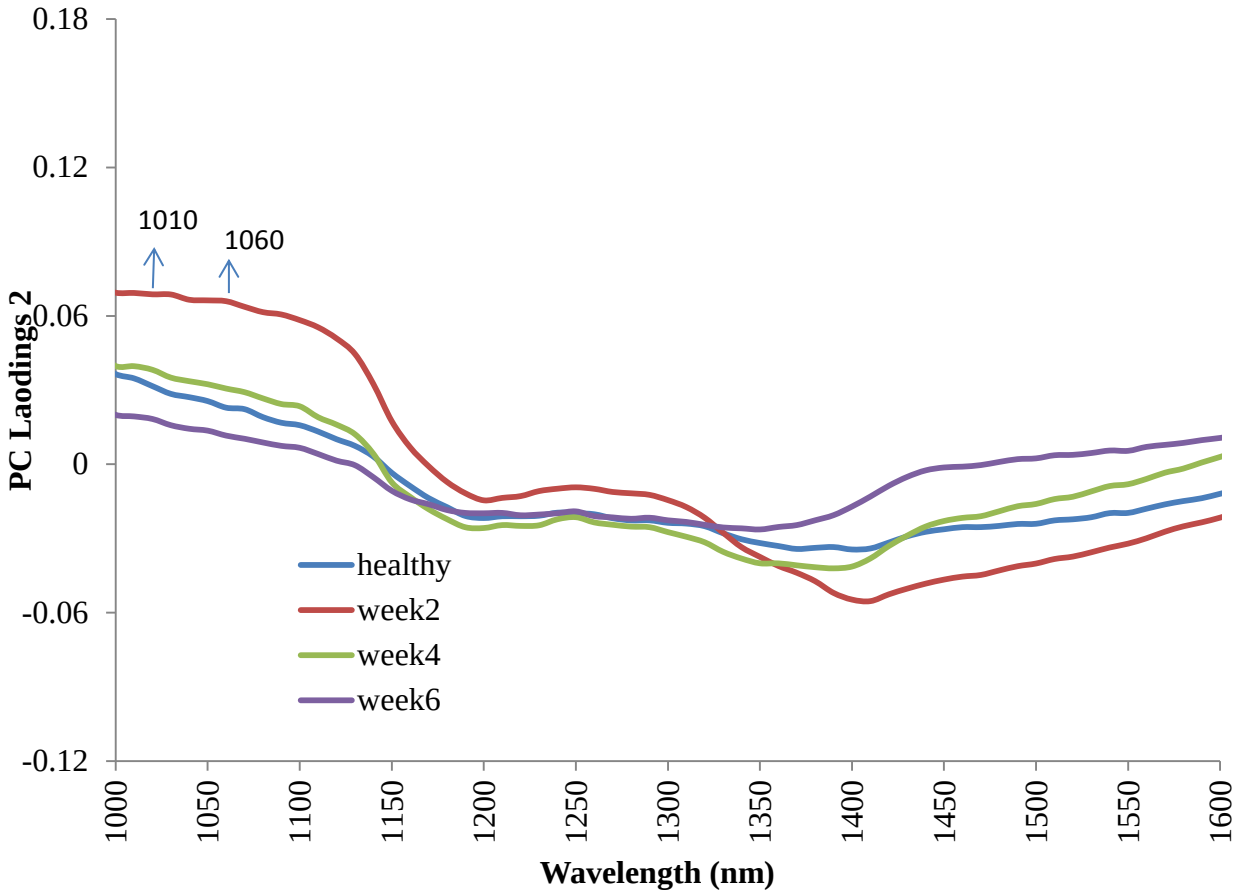

Fig. 4.24 Second PC factor loadings of healthy and (*Penicillium commune*) infected lentils for different weeks of infection stages

Table 4.17 Four-way classification accuracies (mean ± standard deviation, %) of *Penicillium commune* infected lentils using LDA and QDA

	Healthy	Second week	Fourth week	Sixth week
LDA				
Healthy	**100.0±0.0**	0.0±0.0	0.0±0.0	0.0±0.0
Second week	0.0±0.0	**98.2±1.4**	1.7±1.4	0.0±0.0
Fourth week	0.0±0.0	1.0±0.9	**79.1±3.3**	19.7±4.2
Sixth week	0.0±0.0	1.8±0.7	23.2±3.5	**74.8±3.6**
QDA				
Healthy	**100.0±0.0**	0.0±0.0	0.0±0.0	0.0±0.0
Second week	0.6±1.4	**99.2±0.9**	0.0±0.0	0.7±0.9
Fourth week	0.0±0.0	0.0±0.0	**72.5±5.4**	27.4±5.4
Sixth week	0.0±0.0	0.0±0.0	18.6±6.1	**81.3±6.1**

Table 4.18 Two-way classification accuracies (mean ± standard deviation, %) of *Penicillium commune* infected lentils using LDA and QDA

Sample	Statistical Classification using PC1 and PC2 Wavelengths	
	LDA	**QDA**
Healthy	100.0±0.0	100.0±0.0
P.commune week 2	100.0±0.0	100.0±0.0
Healthy	100.0±0.0	100.0±0.0
P.commune week 4	100.0±0.0	100.0±0.0
Healthy	100.0±0.0	100.0±0.0
P.commune week 6	100.0±0.0	100.0±0.0

4.2.4 Pinto bean

Figure 4.25, 4.26 and 4.27 show the mean reflectance, first factor PC loadings and second factor PC loadings, respectively, for healthy and *Penicillium commune* infected pinto bean. Using the features extracted from the five significant wavelengths (Table 4.12), the classification accuracies (mean ± standard deviation, %) using six-way model for healthy and *Penicillium commune* infected pinto bean using LDA and QDA are given in Table 4.19 and using two-way models are given in Table 4.20.

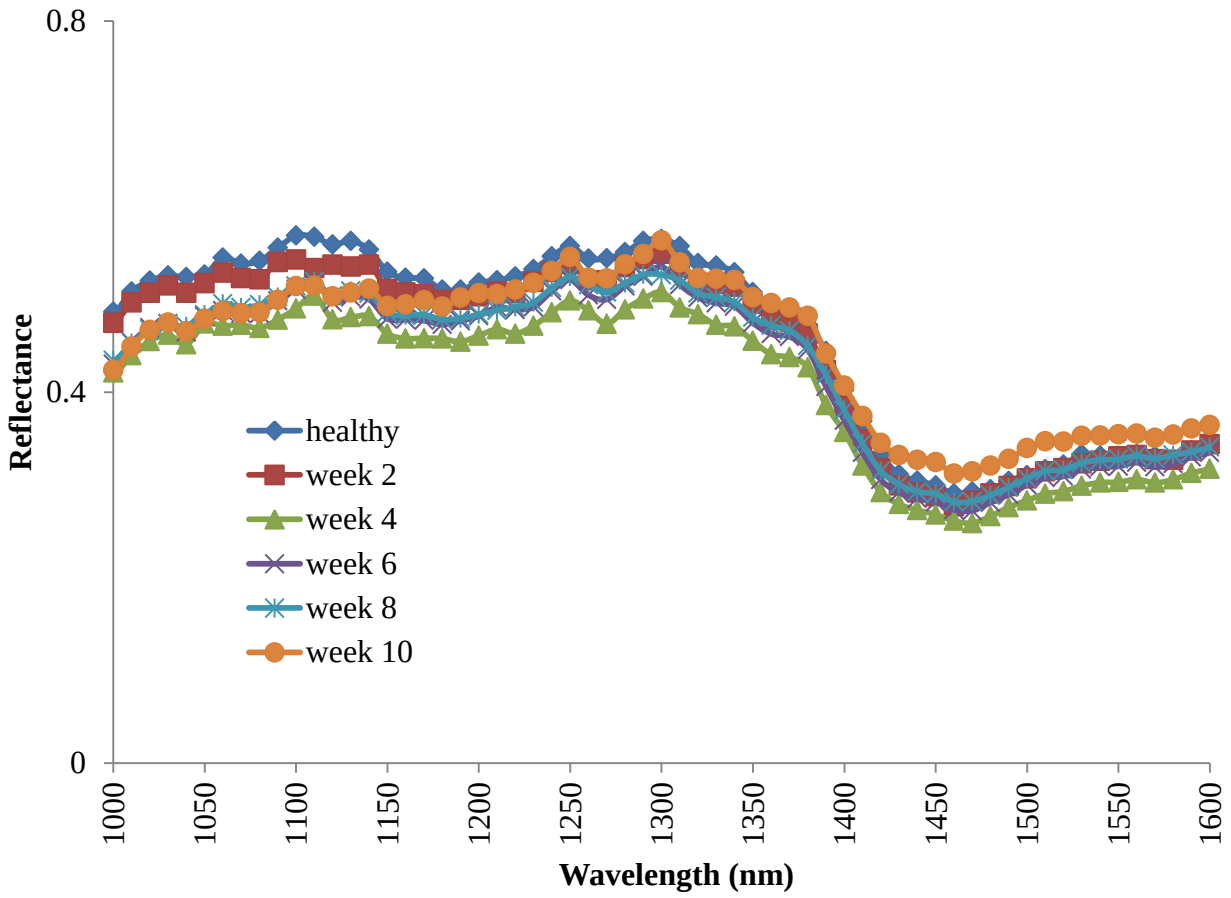

Fig. 4.25 Near-infrared mean reflectance spectra of healthy and infected (*Penicillium commune*) pinto bean.

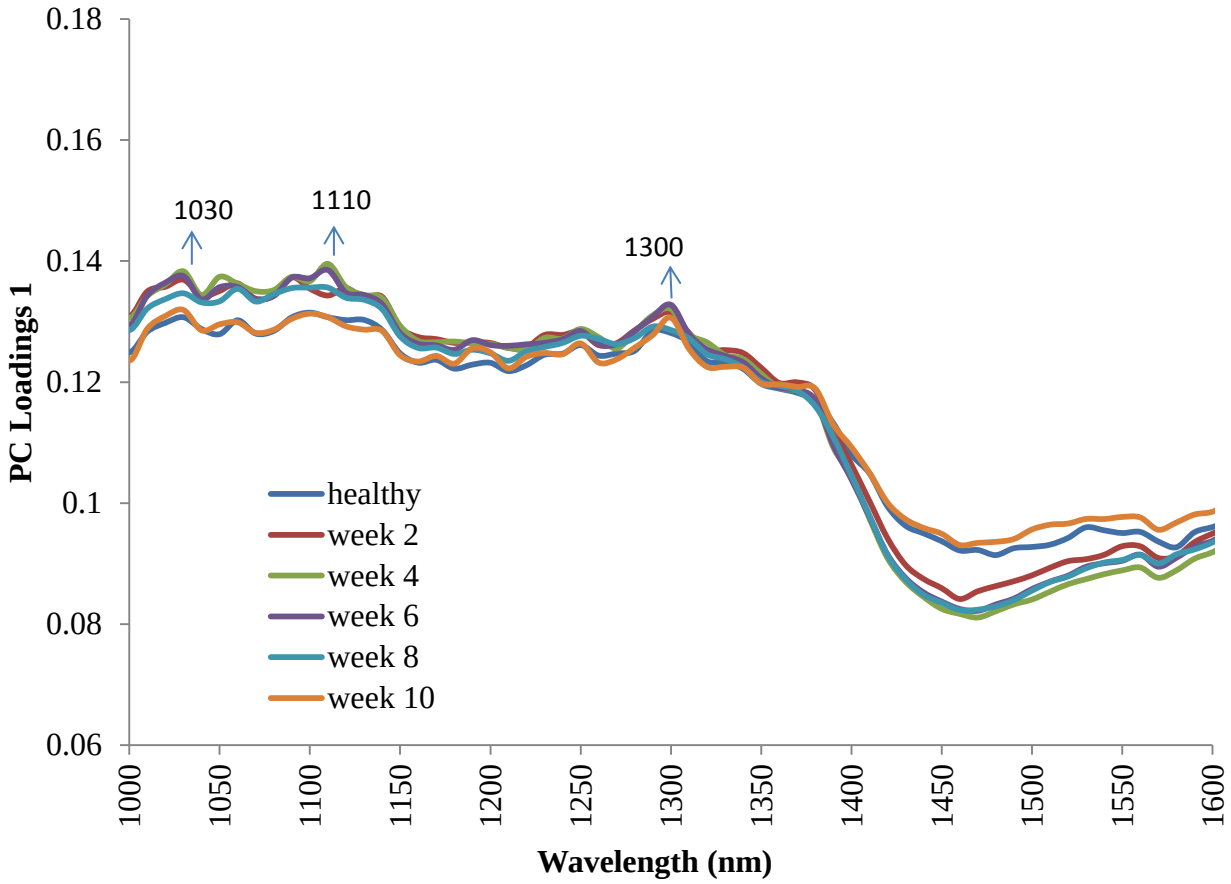

Fig. 4.26 First PC factor loadings of healthy and (*Penicillium commune*) infected pinto bean for different weeks of infection stages

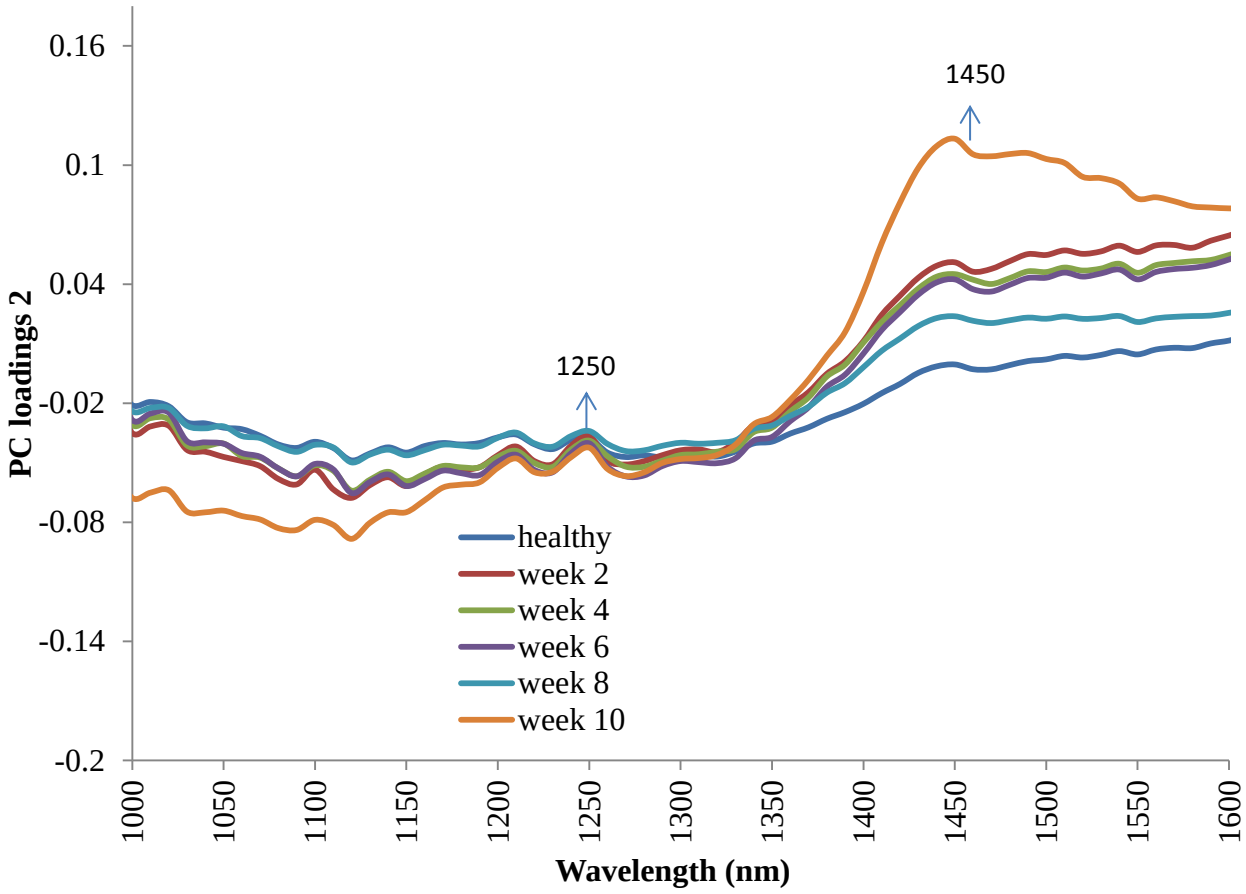

Fig. 4.27 Second PC factor loadings of healthy and (*Penicillium commune*) infected pinto bean for different weeks of infection stages

Table 4.19 Six-way classification accuracies (mean ± standard deviation, %) of *Penicillium commune* infected pinto bean using LDA and QDA

	Healthy	Second week	Fourth week	Sixth week	Eighth week	Tenth Week
LDA						
Healthy	**100.0±0.0**	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0
Second week	0.0±0.0	**97.7±1.5**	1.4±1.5	0.8±1.7	0.0±0.0	0.0±0.0
Fourth week	0.0±0.0	3.8±3.5	**87.4±10.6**	7.0±6.4	1.3±1.4	0.3±0.7
Sixth week	0.0±0.0	0.0±0.0	8.6±8.5	**88.3±10.9**	2.5±1.9	0.4±0.8
Eighth week	0.0±0.0	0.0±0.0	0.2±0.6	3.1±7.1	**93.2±10.8**	3.3±3.2
Tenth week	0.6±0.8	0.0±0.0	0.3±0.7	1.6±3.7	7.8±9.3	**89.4±9.4**
QDA						
Healthy	**98.1±2.2**	0.0±0.0	0.0±0.0	0.0±0.0	1.1±2.4	0.6±0.9
Second week	0.0±0.0	**99.6±0.6**	0.0±0.0	0.3±0.6	0.0±0.0	0.0±0.0
Fourth week	0.0±0.0	0.0±0.0	**85.9±18.0**	10.9±15.0	0.0±0.0	3.6±3.4
Sixth week	0.0±0.0	0.2±0.6	6.4±10.2	**87.5±18.5**	0.0±0.0	5.7±8.6
Eighth week	0.0±0.0	0.0±0.0	0.3±0.7	0.3±0.8	**99.2±0.9**	0.0±0.0
Tenth week	0.0±0.0	0.0±0.0	2.3±2.4	0.6±1.5	0.0±0.0	**97.0±3.6**

Table 4.20 Two-way classification accuracies (mean ± standard deviation, %) of *Penicillium commune* infected pinto bean using LDA and QDA

Sample	Statistical Classification using PC1 and PC2 Wavelengths	
	LDA	**QDA**
Healthy	100.0±0.0	100.0±0.0
P.commune week 2	100.0±0.0	100.0±0.0
Healthy	100.0±0.0	100.0±0.0
P.commune week 4	100.0±0.0	100.0±0.0
Healthy	100.0±0.0	100.0±0.0
P.commune week 6	100.0±0.0	100.0±0.0
Healthy	100.0±0.0	100.0±0.0
P.commune week 8	98.0±2.2	98.2±1.2
Healthy	100.0±0.0	100.0±0.0
P.commune week 10	100.0±0.0	100.0±0.0

4.2.5 Kidney bean

Figure 4.28, 4.29 and 4.30 show the mean reflectance, first factor PC loadings and second factor PC loadings, respectively, for healthy and *Penicillium commune* infected kidney bean. Using the features extracted from the five significant wavelengths (Table 4.12), the classification accuracies (mean ± standard deviation, %) using six-way model for healthy and *Penicillium commune* infected kidney bean using LDA and QDA are given in Table 4.21 and using two-way models are given in Table 4.22.

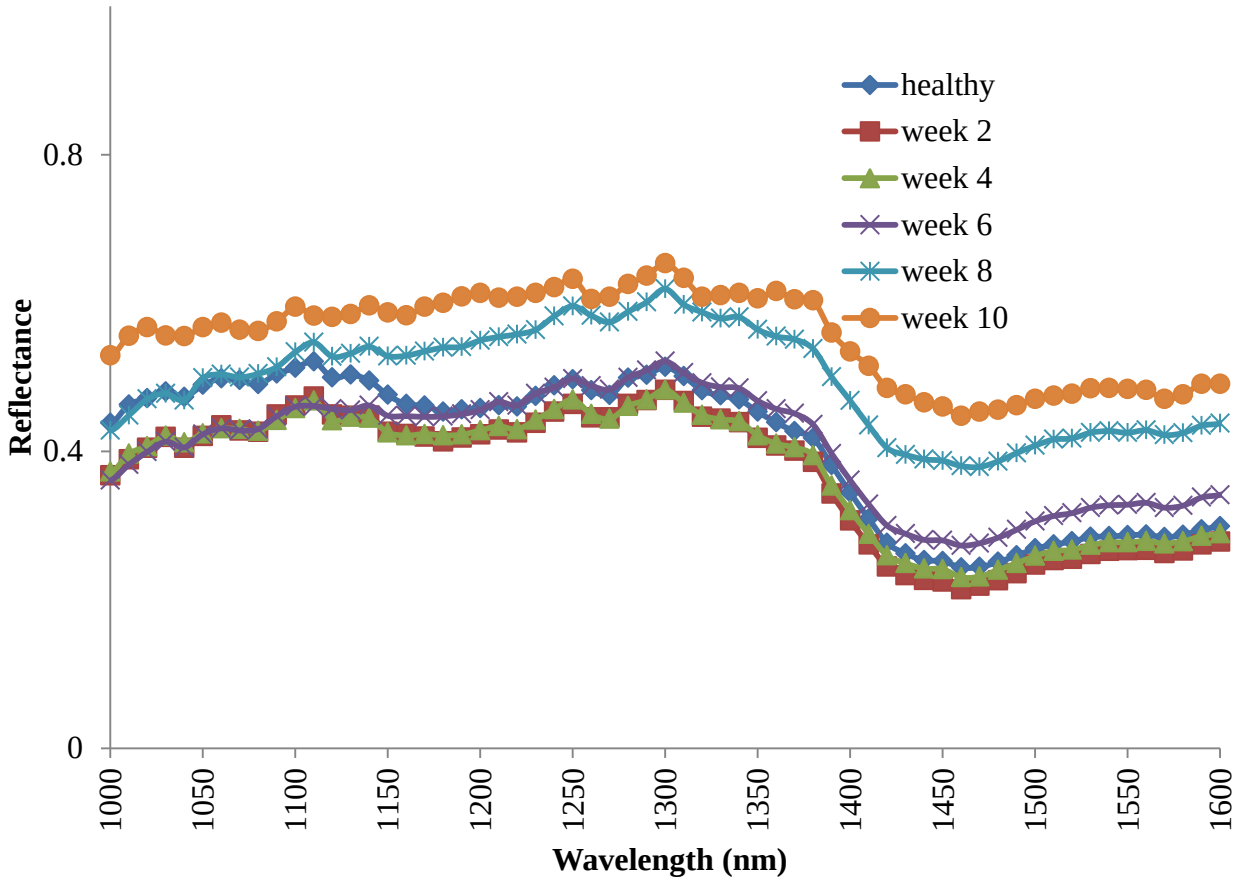

Fig. 4.28 Near-infrared mean reflectance spectra of healthy and infected (*Penicillium commune*) kidney bean

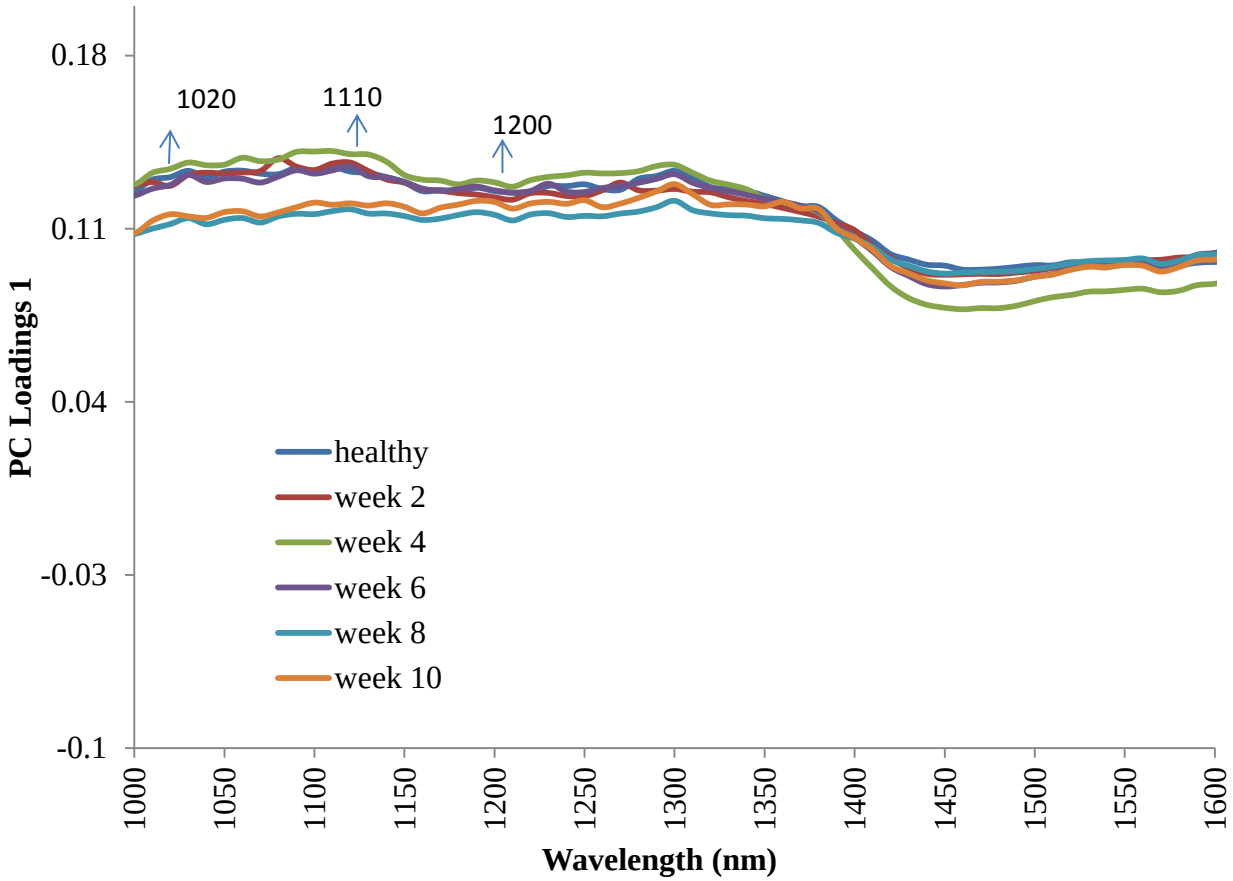

Fig. 4.29 First PC factor loadings of healthy and (*Penicillium commune*) infected kidney bean for different weeks of infection stages

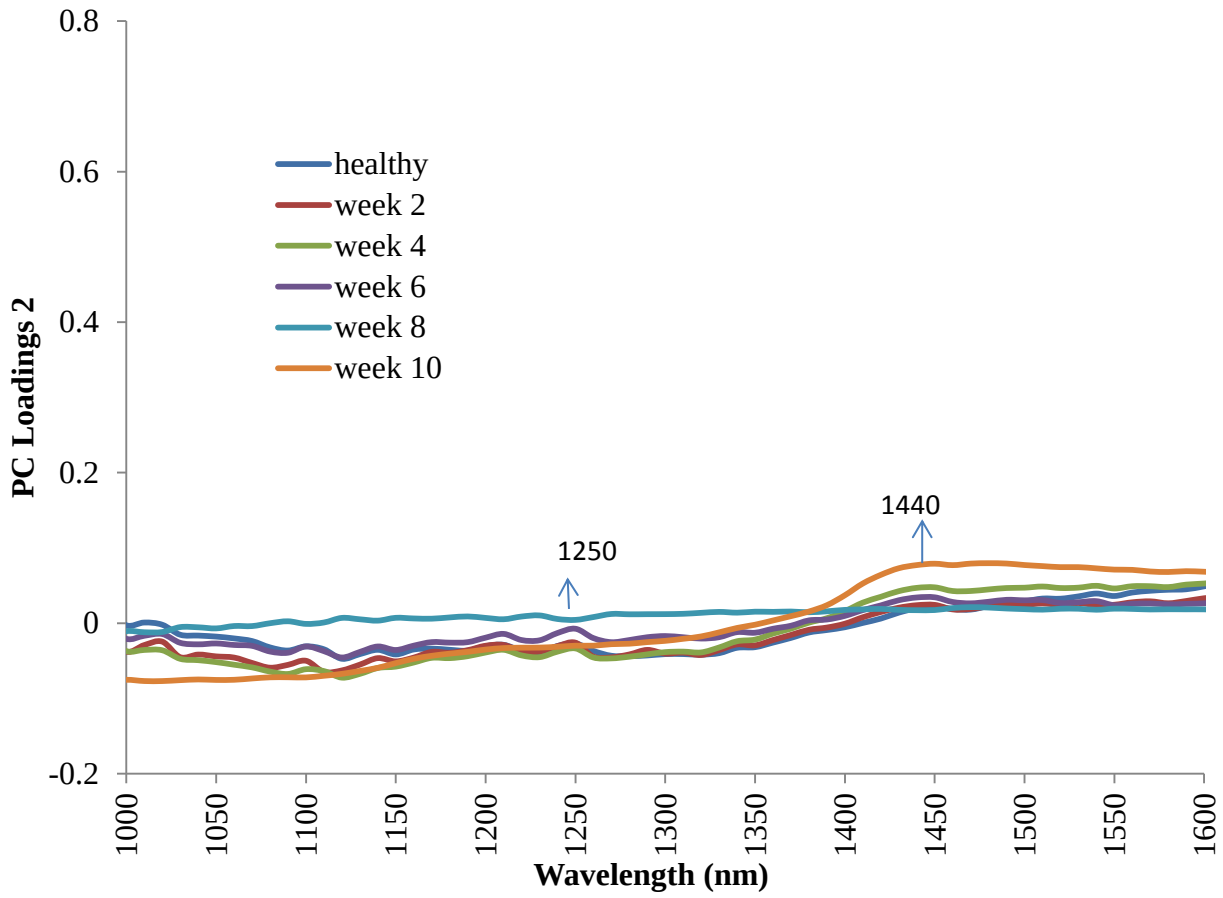

Fig. 4.30 Second PC factor loadings of healthy and (*Penicillium commune*) infected kidney bean for different weeks of infection stage

Table 4.21 Six-way classification accuracies (mean ± standard deviation, %) of *Penicillium commune* infected kidney bean using LDA and QDA

	Healthy	Second week	Fourth week	Sixth week	Eighth week	Tenth week
LDA						
Healthy	**100.0±0.0**	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0	0.0±0.0
Second week	0.0±0.0	**96.9±0.8**	2.2±1.3	0.8±1.7	0.0±0.0	0.0±0.0
Fourth week	0.9±2.1	5.0±3.1	**83.2±9.2**	9.7±5.5	0.6±0.8	0.3±0.7
Sixth week	0.0±0.0	0.0±0.0	8.3±5.2	**86.4±8.2**	4.9±6.8	0.3±0.6
Eighth week	0.0±0.0	0.0±0.0	0.0±0.0	2.4±5.3	**93.1±8.5**	4.4±3.2
Tenth week	0.6±0.8	0.0±0.0	0.3±0.7	0.9±2.1	13.4±9.2	**84.5±8.4**
QDA						
Healthy	**98.1±2.2**	0.0±0.0	0.0±0.0	0.0±0.0	1.1±2.4	0.7±0.9
Second week	0.2±0.6	**97.5±3.7**	0.5±1.3	1.2±1.2	0.0±0.0	0.2±0.6
Fourth week	0.0±0.0	0.0±0.0	**86.7±16.6**	10.5±14.7	0.0±0.0	2.6±3.0
Sixth week	0.0±0.0	0.2±0.6	7.5±10.1	**87.6±18.2**	0.6±1.5	3.8±8.6
Eighth week	0.0±0.0	0.0±0.0	0.3±0.7	0.7±1.0	**95.7±6.4**	3.1±6.9
Tenth week	0.0±0.0	0.0±0.0	0.9±2.0	0.0±0.0	1.7±3.8	**97.3±3.8**

Table 4.22 Two-way classification accuracies (mean ± standard deviation, %) of *Penicillium commune* infected kidney bean using LDA and QDA

Sample	Statistical Classification using PC1 and PC2 Wavelengths	
	LDA	**QDA**
Healthy	100.0±0.0	100.0±0.0
A. flavus week 2	95.0±1.0	92.0±2.2
Healthy	100.0±0.0	100.0±0.0
A. flavus week 4	100.0±0.0	98.4±3.4
Healthy	100.0±0.0	100.0±0.0
A. flavus week 6	98.4±2.0	100.0±0.0
Healthy	98.0±3.0	100.0±0.0
A. flavus week 8	100.0±0.0	98.0±2.2
Healthy	100.0±0.0	100.0±0.0
A. flavus week 10	90.0±0.0	97.0±3.3

4.3 Classification for healthy vs similar weeks of infections by *A. flavus* and *Penicillium commune*.

When similar weeks of infections by two fungal species were classified against healthy samples, the accuracies were low for all pulse types. For example, the mean classification accuracy of healthy chick peas, A. *flavus* infected week 2, and *Penicillium commune* infected week 2 was 50.0±2.2% and the accuracies for weeks 4, 6, 8 and 10 were 47.4±1.4%, 40.9±2.4%, 30.3±6.4% and 30.2±9.4%, respectively.

Similarly, the mean classification accuracy of healthy green peas, *A. flavus* infected week 2 and *Penicillium commune* infected week 2 was 49.0±2.2% and the accuracies for weeks 4, 6, 8 and 10 were 46.4±2.4%, 45.4±4.4%, 35.4±4.4% and 29.4±7.4%, respectively.

Similarly, the mean classification accuracies of healthy lentils vs *A. flavus* infected week 2 and *Penicillium commune* infected week 2 was 51.0±2.3% and the accuracies for weeks 4, 6, 8 and 10 were 45.4±7.4%, 42.4±4.4%, 37.4±9.4% and 32.4±10.4%, respectively.

Similarly, the mean classification accuracies of healthy pinto bean vs *A. flavus* infected week 2 and *Penicillium commune* infected week 2 was 49.9±2.6% and the accuracies for weeks 4, 6, 8 and 10 were 48.4±1.4%, 44.4±2.4%, 30.4±6.4% and 25.4±9.4%, respectively.

Similarly, the mean classification accuracies of healthy red kidney vs *A. flavus* infected week 2 and *Penicillium commune* infected week 2 was 53.0±2.6% and the accuracies for weeks 4, 6, 8 and 10 were 42.4±6.4%, 40.4±2.4%, 35.4±7.4% and 30.4±3.4%, respectively.

These results show that fungal species cannot be identified very accurately but fungal infection by either species can be detected with high accur

5. Discussion

Table 5.1 Application of exploratory data analysis techniques for hyperspectral data analysis in food and agro products

Product	Factor assessed	Imaging Wavelength (nm)	Significant Wavelength (nm)	Statistics	Reference
Wheat	α-amylase activity	1235-2450	1350,1450	PCA	Xing et al. (2011)
Cereals	Viability	1000-2500	-	PCA	McGoverin et al. (2011)
Mung Bean	Insect damage	1000-1600	1100, 1290 and 1450	PCA	Kaliramesh et al. (2013)
Spinach	*E. Coli*	400-1000	-	PCA	Siripatrawan et al. (2011)
Wheat	Classes	960-1700	1200, 1700	LDA, QDA, ANN	Mahesh et al. (2008)
Wheat	Classes	100-1600	-	SVM	Huang et al. (2007)
Wheat	Wavelet features	960-1700	960, 990, 1090, 1210, 1450.	LDA, QDA, ANN	Choudhary et al. (2009)
Wheat	Midge damage	700-1100	870, 890	LDA, QDA	Singh et al. (2010)
Wheat	Classes, moisture content	960-1700	1060, 1090, 1340 and 1450	LDA, QDA	Mahesh et al. (2011)
Wheat	Insect damage	1000-1600	1101, 1305	LDA, QDA	Singh et al. (2009)
Mung bean	Insect damage	1000-1600	1100, 1290 and 1450	LDA, QDA	Kaliramesh et al. (2013)
Wheat	*Fusarium*	400-1000	500, 560, 682 and 927	PCA	Bauriegel et al. (2011)

Maize	Fungal growth	1000-2498	1900, 2136	PCA	Williams et al. (2012)
Maize	*Fusarium*	960-2662	1450, 1960 and 2100	PCA	Williams et al. (2010)
Maize	Toxigenic Fungi	400-1000	410, 535 and 945	PCA	Del Fiore et al. (2010)
Wheat	Fungal Detection	1000-1600	1350, 1450 and 1550	PCA	Singh et al. (2007)
Canola	Fungal detection	960-1700	960, 1300 and 1420	PCA	SenthilKumar et al. (2012)
Wheat	Fungal detection	960-1700	960, 1060, 1330, 1390,1480 and 1680	PCA	Wang et al. (1999)
Date Fruits	Fungal infection	960-1700	1120, 1300, 1610 and 1650	PCA	Teena et al. (2014)

Table 5.1 shows the significant wavelengths identified from the hyperspectral data analyses for assessment of food and agro products. These significant wavelengths are compared with wavelengths identified in this study from fungal infection by *A. Flavus* (Table 4.1) and *P. Commune* (Table 4.12).Senthilkumar et al. (2012) reported that the 1300 nm wavelength could be attributed to fungal infection and 960 and 1420 nm wavelengths correspond to water content in pulses. Osborne (2006) reported that the wavelength region 1100-1300 nm corresponds to carbon-hydrogen first and second overtones and carbon-hydrogen combination bands and the significance can be associated with the absorption by starch molecules. Wang et al. (1999) pointed out that the wavelengths 960, 1060, 1330, 1390, 1480 and 1680 nm correspond to kernel

hardness and 1390 nm to oil content in wheat. Kaliramesh et al. (2013) pointed out that the wavelengths 1100, 1290 and 1450 nm were significant for the classification of fungal infected mung bean pulses. Teena et al. (2014) reported that the wavelengths 1120, 1300, 1610 and 1650 nm were prominent for the classification of *A. flavus* infected date fruits.

Williams et al. (2010) reported that the wavelengths 1450, 1960 and 2100 nm wavelengths were useful for *fusarium* detection on maize kernels. Singh et al. (2007) found out that the 1450 and 1550 nm were significant for fungal detection in wheat. For wheat classes classification and moisture content prediction Mahesh et al. (2011) predicted that the wavelengths 1340 and 1450 nm were significant.

In this study, five identified wavelength regions (Tables 4.1, 4.12) are in the similar ranges for both *A. flavus* and *P. commune* infected pulses. The wavelengths 1110, 1250, 1300, 1440 and 1550 nm were mostly common to both types of fungal infection. Hence a machine vision system can be designed on an industrial scale where the device can be programmed to capture images only at these significant wavelengths reducing the capturing time and easily removing the infected pulses.

Using the combined features of six statistical features (maximum, minimum, mean, median, variance and standard deviation) extracted from the significant wavelengths and 50 histogram features (10 histogram features at each wavelength) LDA and QDA classification models were developed to identify infections at different stages by both fungi for all five pulse types. The reason for combining these features was that the histogram features alone did not classify accurately as it produced non-positive covariance matrix which might be due to linear dependency of one or more input features.

The reason why two-way classification model produced better results than six-way model is that in two-way we just use healthy and one infection whereas in six-way model we use healthy and all the different infection stages for classification which could misclassify the results on a larger scale. In six-way models, most of the misclassifications were to nearest levels of infections where differences probably overlap. For example, 8^{th} week infection level got misclassified to 6^{th} week infection level.

6. Conclusion

The results from this study indicate that hyperspectral imaging is a promising technique to detect fungal infection in early stages in stored pulses. In total, 1800 (300 healthy and 1500 for a period of 10 weeks, 300 every two weeks) seeds from each pulse type with two different fungal infection namely *Aspergillus flavus* and *Penicillium commune* were imaged using the near infrared hyperspectral imaging system.

For *A. flavus* infected pulses, the six-way classification of healthy pulses using LDA showed close to 100.0±0.0% classification accuracy for all the five types of pulses but the classification accuracies varied for different weeks (weeks 2, 4, 6, 8 and 10) of infections from 77.4±6.4 to 96.2±2.1% for chickpeas, 99.3±0.9 to 98.2±1.6% for green peas, 71.6±5.3 to 100.0±0.0% for lentils, 88.0±10.5 to 97.5±2.1% for pinto bean and 95.5±8.9 to 99.0±1.4% for kidney bean, respectively.

The QDA showed 100.0±0.0% classification accuracy of healthy pulses for all the five types of pulses but the classification accuracies varied for different weeks (weeks 2, 4, 6, 8 and 10) of infections from 71.1±4.3 to 89.9±5.0% for chickpeas, 98.9±0.9 to 99.6±0.6% for green peas, 69.1±18.6 to 100.0±0.0% for lentils, 75.1±13.6 to 94.5±5.8% for pinto bean and 87.1±16.9 to 97.0±3.6% for kidney bean, respectively.

For two-way classification of *A. flavus* infected pulses, both LDA and QDA classifiers almost showed 100.0±0.0% classification accuracies for all the healthy vs. different weeks of infection (weeks 2, 4, 6, 8 and 10).

For *P. commune* infected pulses, the four way (after sixth week the chick peas and lentils pulse was badly infected and not suitable for imaging) classification of healthy pulses using LDA showed 100.0±0.0% classification accuracy for chick peas and lentils but the classification accuracies varied for different weeks (weeks 2, 4, 6) from 86.6±4.2 to 99.3±10.9% for chickpeas, 74.8±3.6 to 98.2±1.4% for lentils, respectively.

The six-way classification for *P. commune* using LDA for healthy pulses showed 96.8±2.0% classification accuracy for green peas and 100.0±0.0% for both pinto bean and kidney bean but the classification accuracies varied for different weeks (weeks 2, 4, 6, 8 and 10) from 70.9±6.3 to 97.6±1.0% for green peas, 87.4±10.6 to 97.7±1.5% for pinto bean and 83.2±9.2 to 96.9±0.8% for kidney beans, respectively. For QDA the four way classification of healthy pulses using LDA showed 100.0±0.0% classification accuracy for chick peas and lentils but the classification accuracies varied for different weeks (weeks 2,4,6) from 85.9±6.8 to 95.8±3.4% for chickpeas, 72.5±8.4 to 99.2±0.9% for lentils, respectively. The six-way classification using QDA for healthy pulses showed 95.9±2.0% classification accuracy for green peas and 100.0±0.0% for both pinto bean and kidney bean but the classification accuracies varied for different weeks (weeks 2, 4, 6, 8 and 10) from 66.1±4.9 to 94.5±6.0% for green peas, 85.9±18.0 to 99.6±0.6% for pinto bean and 86.7±16.6 to 97.5±3.7% for kidney bean, respectively. For two-way classification of *P. commune* infected pulses, both LDA and QDA classifiers almost showed 100.0±0.0% classification accuracies for all the healthy vs. different weeks of infection (weeks 2, 4, 6, 8 and 10). Hence both LDA and QDA classifiers prove to be a promising technique for detecting the fungal infection in pulses.

The feasibility of hyperspectral imaging to detect fungal infection in pulses using different fungal species such as *Fusarium* spp. should be studied in the future. With this work we were

able to demonstrate the hyperspectral imaging has the potential for detection of fungal infections in pulses. It could be possible to use hyperspectral imaging to detect fungal infection in different seeds and food materials. In future, combinations of insect and fungal infestation levels in pulses should be studied using hyperspectral imaging.

7. References

Abbott, J.A. 1999. Quality measurement of fruits and vegetables. *Journal of Postharvest Biology and Technology* 15(3): 207-225.

Ariana D.P., R. Lu and D.E. Guyer. 2006. Near-infrared hyperspectral reflectance imaging for detection of bruises on pickling cucumbers. *Journal of Computer and Electronics in Agriculture* 53(1): 60-70.

Bauriegel E., A. Giebel, M. Geyer, U. Schmidt and W.B. Herppich 2011. Early detection of Fusarium infection in wheat using hyperspectral imaging. *Computer and Electronics in Agriculture* 75(2): 304-312.

Chelladurai V., S. Kaliramesh and D.S. Jayas. 2012. Detection of *Callosobruchus maculatus* (F.) infestation in mung bean (*Vigna radiata*) using thermal imaging technique. *Food and Bioprocessing* CSBE 12121.

Choudhary R., S. Mahesh, J. Paliwal and D.S. Jayas 2009. Identification of wavelet features from near infrared hyperspectral images of bulk samples. *Journal of Biosystems Engineering* 102(2): 115-127.

Daoyi C., J. Huang and T.J. Jackson. 2005. Vegetation water content estimation for corn and soybeans using spectral indices derived from MODIS near- and short-wave infrared bands. *Remote Sensing of Environment* 98:2-3.

Del Fiore, A., M. Reverberi, A. Ricelli, F. Pinzari, S. Serranti, A.A. Fabbri, G. Bonifazi and C. Fanelli. 2010. Early detection of toxigenic fungi on maize by hyperspectral imaging analysis. *International Journal of Food Microbiology* 144(1): 64-71.

Delwiche S.R., M.S. Kim and Y. Dong. 2011. *Fusarium* damage assessment in wheat kernels by Vis/NIR hyperspectral imaging. *Journal of Sensory and Instrumentation Food Quality* 5: 63-71.

Dolmatova, L., C. Ruckebusch, N. Dupuy, J. Huvenne and P. Legrand. 1998. Identification of modified starches using infrared spectroscopy and artificial neural network processing. *Applied Spectroscopy* 52(3): 329-338.

Einsele, H., H. Hebart, G. Roller, J. Loffler, I. Rothenhofer, C.A. Muller, R.A. Bowden, J.V. Burik, D. Engelhard, L. Kanz and U. Schumache. 1997. Detection and identification of fungal pathogens in blood by using molecular probes. *Journal of Clinical Microbiology* 35(6): 1353-1360.

ElMasry, G., N. Wang, C. Vigneault, J. Qiao and A. ElSayed. 2008. Early detection of apple bruises on different background colors using hyperspectral imaging. *LWT- Food Science and Technology* 41(2): 337-345.

FAO. 2014. FAO Cereal Supply and Demand Brief: World Food Situation. Food and Agriculture Organization of the United Nations. http://www.fao.org/worldfoodsituation/csdb/en/. Accessed 17 Apr 2014.

Favro LD., X. Han, Z. Ouyang, G. Sun, H. Sui and R.L. Thomas. 2000. Infrared imaging of defects heated by a sonic pulse. *American Institute of Physics* 71:2418.

Geladi P. and H. Grahn. 1996. Multivariate Image Analysis. John Wiley and Sons, Chichester, UK.

Gowen, A.A., C.P. O`Donnell, P.J. Cullen, G. Downey and J.M. Frias. 2007. Hyperspectral imaging - an emerging process analytical tool for food quality and safety control. *Trends in Food Science and Technology* 18(12): 590-598.

Guyer, D. and X. Yang. 2000. Use of genetic artificial neural networks and spectral imaging for defect detection on cherries. *Computers and Electronics in Agriculture* 29(3): 179-194.

Holley, L. and M. Karplus. 1989. Protein secondary structure prediction with a neural network. *Proceedings of the National Academy of Sciences of the United States of America* 86(1): 152-156.

Huang M., X. Wan, M. Zhang and Q. Zhu. 2013. Detection of insect damaged vegetable soybeans using hyperspectral transmittance image. *Journal of Food Engineering* 116(1): 45-49.

Huang W., D.W. Lamb, Z.Niu, Y. Zhang, L. Liu and J. Wang 2007. Identifiaction of yellow rust in wheat using in-situ spectral reflectance measurements and airborne hyperspectral imaging. *Journal of precision agriculture* 8: 187-197.

Johnson, R.W., D.W. Wilson, W. Fink, M. Humayun and G. Bearman. 2007. Snapshot hyperspectral imaging in Ophthalmology. *Journal of Biomedical Optics* 12(1):014036.

Joyce B., F. Zare and A. Pletch. 2010. Pulse proteins: processing, characterization, functional properties and applications in food and feed. *Journal of Food Research International* 43(2): 414-431.

Kaliramesh S., V. Chelladurai, D.S. Jayas and K. Alagusundaram. 2014. Determination of main constituents in green gram using near infrared hyperspectral imaging. *Journal of Agricultural Engineering* 51.

Kaliramesh, S., V. Chelladrai, D.S. Jayas, K. Alagusundaram, N.D.G. White and P.G. Fields. 2013. Detection of infestation by *Callosobruchus maculatus* in mungbean using near-infrared hyperspectral imaging. *Journal of Stored Products Research* 52: 107-111.

Kamruzzaman M., G. Elmasry, D.W. Sun and P. Allen. 2011. Application of NIR hyperspectral imaging for discrimination of lamb muscles. *Journal of Food Engineering* 104(3): 332-340.

Khoobehi B., J.M. Beach and H. Kawano. 2004. Hyperspectral imaging for measurement of oxygen saturation in the optic nerve head. *Journal of Investigative Ophthalmology and Visual Science* 45: 1464-1472.

Mahesh S., A. Manickavasagan, D.S. Jayas, J. Paliwal and N.D.G. White. 2008. Feasibility of near-infrared hyperspectral imaging to differentiate Canadian wheat classes. *Biosystems Engineering* 101(1): 50-57.

Mahesh S., D.S. Jayas, J. Plaiwal and N.D.G. White 2011. Identification of wheat classes at different moisture levels using near-infrared hyperspectral images of bulk samples. *Sensing and instrumentation for Food quality and safety* 5(1): 1-9.

McGoverin C.M., P. Engelbrecht, P. Geladi and M. Manley 2011. Characterisation of non-viable whole barley, wheat and sorghum grains using near-infrared hyperspectral data and chemometrics. *Journal of Analytical and Bioanalytical Chemistry* 401: 5291.

Menlik, T., M. Ozdemir and V. Kirmaci. 2010. Determination of freeze-drying behaviours of apples by artificial neural network. *Expert Systems with Applications* 37(12): 7669-7677.

Monteiro S.T., Y. Minekawa, Y. Kosugi, T. Akazawa and K. Oda. 2007. Prediction of sweetness and amino acid content in soybean crops from hyperspectral imagery. *ISPRS Journal of Photogrammetry and Remote Sensing* 62(1): 2-12.

Osborne. 2006. NIR spectroscopy in food analysis. *Encyclopedia of Analytical Chemistry.* John Wiley and Sons Ltd., Chichester, UK.

Park B., K.C. Lawrence, W.R. Windham and D.P. Smith. 2006. Performance of hyperspectral imaging system for poultry surface fecal contaminant detection. *Journal of Food Engineering* 75(3): 340-348.

Qiao J., M.O. Ngadi, N. Wang, C. Gariepy and S.O. Prasher. 2007. Pork quality and marble level assessment using a hyperspectral imaging system. *Journal of Food Engineering* 83(1): 10-16.

Saleem U., M. Schlerf, A.K. Skidmore and C. Hecker. 2012. Identifying plant species using mid-wave infrared and thermal infrared emissivity spectra. *Remote Sensing of Environment* 118: 95-102.

Schnurer, J., J. Olsson and T. Borjesson. 1999. Fungal volatiles as indicators of food and feeds spoilage. *Journal of Fungal Genetics and Biology* 23(2-3): 209-217.

Senthilkumar T., C.B. Singh, D.S. Jayas and N.D.G. White. 2012. Detection of fungal infection in canola using Near-Infrared Hyperspectral Imaging. *Journal of Agricultural Engineering* 49(1): 2012.

Singh C.B., D.S. Jayas, J. Paliwal and N.D.G. White. 2010. Detection of midge-damaged wheat kernels using short-wave near-infrared hyperspectral and digital colour imaging. *Journal of Biosystems Engineering* 105(3): 380-387.

Singh, C.B., D.S. Jayas, J. Paliwal and N.D.G. White. 2007. Fungal detection in wheat using near-infrared hyperspectral imaging system. *Transactions of the ASABE* 50(6): 2171-2176.

Singh, C.B., D.S. Jayas, J. Paliwal and N.D.G. White. 2009. Detection of insect-damaged wheat kernels using near-infrared hyperspectral imaging. *Journal of Stored Products Research* 45(3): 151-158.

Siripatrawan U., Y. Makino, Y. Kawagoe and S. Oshita. 2011. Rapid detection of *Escherichia coli* contamination in packaged fresh spinach using hyperspectral imaging. *Journal of Talanta* 85(1): 276-281.

Statistics Canada (2011). Census of Agriculture: August. Canadian Grain Commission Industry Services. Statistics Canada. http://www.statcan.gc.ca/pub/96-325-x/2014001/article/14041-eng.htm#wb-tphp. Accessed 06 Oct 2015.

Tamara L.T. and S.N. Thennadil. 2000. Optical properties of human skin in the near infrared wavelength range of 1000 to 2200 nm. *Journal of Biomedical Optics* 6(2): 167-176.

Teena M.A., A. Manickavasagan, L. Ravikanth and D.S. Jayas. 2014. Near infrared (NIR) hyperspectral imaging to classify fungal infected date fruits. *Journal of Stored Products Research* 59: 306-313.

Vermeulen Ph., J.A. Fernandez, H.P. Van, P. Dardenne and V. Baeten. 2012. Online detection and quantification of ergot bodies in cereals using near infrared hyperspectral imaging. *Journal of Food Additives and Contaminants* 29(2): 232-240.

Wang D., F.E. Dowell and R.E. Lacey. 1999. Single wheat kernel color classification using neural networks. *Transactions of the ASAE* 42: 232-240.

Wang L., D. Liu, H. Pu, D.W. Sun, W. Gao and Z. Xiong. 2015. Use of hyperspectral imaging to discriminate the variety and quality of rice. *Journal of Food Analysis Methods* 8: 515-523.

Williams P., M. Manley, G. Fox and P. Geladi 2010. Indirect detection of Fusarium verticillioides in maize (*Zea maize* L.) kernels by NIR hyperspectral imaging. *Journal of Near Infrared Spectroscopy* 18(1): 49-58.

Williams P.J., P. Geladi, T.J. Britz and M. Manley 2012. Investigation of fungal development in maize kernels using NIR hyperspectral imaging and multivariate data analysis. *Journal of cereal science* 55(3): 272-278.

Williams, R.H., E. Ward and H.A. McCartney. 2001. Methods for integrated air sampling and DNA analysis for detection of airborne and fungal spores. *Journal of Applied and Environmental Microbiology* 67(6): 2453-2459.

Xing J. and J.D. Baerdemaeker. 2005. Bruise detection on 'Jonagold' apples using hyperspectral imaging. *Journal of Postharvest Biology and Technology* 37(2): 152-162.

Xing J., S. Symons, D. Hatcher and M. Shahin. 2011. Comparison of short-wavelength infrared (SWIR) hyperspectral imaging system with an FT-NIR spectrophotometer for predicting alpha-amylase activities in individual Canadian western red spring (CWRS) wheat kernels. *Journal of Biosystems Engineering* 108(4): 303-310.

Zhang X and Y. He. 2013. Rapid estimation of seed yield using hyperspectral images of oilseed rape leaves. *Journal of Industrial Crops and Products* 42: 416-420.

Zhang X., F. Liu and X. Gong. 2013. Detecting macronutrients content and distribution in oilseed rape leaves based on hyperspectral imaging. *Biosystems Engineering* 115(1): 56-65.

Zuzak, K.J., M.D. Schaeberle, E.N. Lewis and I.W. Levin. 2002. Visible reflectance hyperspectral imaging: Characterization of a noninvasive, in vivo system for determining tissue perfusion. *Journal of American Chemical Society* 74(9): 2021-2028.