

Fourier Transform Infrared Microspectroscopy, with Attenuated Total Reflectance, as a
Detection Method for Fusarium in Whole Wheat Kernels

by

Karen Starr

A Thesis submitted to the Faculty of Graduate Studies of
The University of Manitoba
in partial fulfilment of the requirements of the degree of

MASTER OF SCIENCE

Department of Biosystems Engineering
University of Manitoba
Winnipeg

Copyright © 2012 by Karen Starr

ABSTRACT

The purpose of this study was to investigate the feasibility of using Fourier transform mid-infrared (FT-IR) microspectroscopy, with attenuated total reflectance (ATR), to detect the presence of *Fusarium* spp., classify differences in ability to detect *Fusarium* spp. among the Canadian soil zones, determine whether *Fusarium* spp. can serve as an indicator for the concentration of deoxynivalenol (DON), and develop an equation that can predict DON within wheat kernels. Canadian Western Red Spring wheat kernels were pressed against an ATR crystal to obtain spectra. Five chemical functional groups were investigated. The ability of FT-IR-ATR microspectroscopy to detect the presence of *Fusarium* spp. was confirmed. The chemical composition, and the ability to identify *Fusarium* spp., differed in kernels grown among the soil zones. Several methods to use *Fusarium* spp. as an indicator for the concentration of DON content within the wheat kernels were attempted and found to be unsuccessful.

ACKNOWLEDGEMENTS

I'd like to thank my advisor, Dr. Jitendra Paliwal, for helping me fulfill an academic dream of mine to undertake graduate studies in engineering. His support and knowledge through this process was fundamental to my successful completion of this program. I would also like to thank Dr. Jason Morrison and Dr. Kathleen Gough for serving on my advisory committee and continually providing assistance during the course of this project.

I am thankful to the staff of the grain quality department at Cargill Ltd. for supplying an endless amount of wheat samples. Additionally, I'd like to thank Dr. Curtis Rempel and his staff at the Richardson Food Center for Functional Foods and Nutraceuticals for supplying testing services that were completely unique to my request. Furthermore, I would like to thank Richard Wiens for training me on the instrumentation used in this project. The project was supported by the Agricultural and Rural Development Initiative and the University of Manitoba Graduate Fellowships.

Finally, I would like to thank my family and friends for the immense support they have given me through this process. I would not be where I am without their undying encouragement and for that, I am beyond thankful.

TABLE OF CONTENTS

ABSTRACT.....	i
ACKNOWLEDGEMENTS.....	ii
LIST OF TABLES.....	vii
LIST OF FIGURES.....	viii
1.INTRODUCTION.....	1
1.1 Canadian Grain Industry: Present and Future.....	1
1.2 Fusarium Fungi: The Need for an Objective Measurement.....	3
1.3 Fourier Transform Infrared Microspectroscopy as a Viable Technology for Grain Quality Assessment.....	4
1.4 Objectives of Research.....	6
2.LITERATURE REVIEW.....	9
2.1 History and Background of Fourier Transform Infrared Spectroscopy.....	9
2.2.1 Vibrational-Rotational Energy of Molecules:.....	13
2.2.2 Molecular Degrees of Freedom and Spectroscopic Selection Rule.....	17
2.2.3 Anharmonic Model:.....	19
2.2.4 Fourier Transforms.....	22
2.3 Measurement Modes:.....	27
2.3.1 Interaction with Matter:.....	27

2.3.1 Focal Plane Array Detectors	29
2.3.2 Attenuated Total Reflectance.....	30
2.4 Instrumentation:	35
2.4.1 Overview of FT-IR Spectrometers.....	36
2.4.2 Interferometer:	37
2.4.3 Generation of an Interferogram.....	41
2.5 Chemometrics	43
2.5.1 Pre-processing Methods:.....	44
2.5.2 Quantitative Analysis.....	51
2.6 FT-IR for Fusarium Detection	57
2.6.1 Fusarium Head Blight	57
2.6.2 Current Applications	63
3. METHODOLOGY	72
3.1 System Design	72
3.1.1 Selection of FT-IR	72
3.1.2 Selection of Measurement Mode	73
3.1.3 Design of Kernel Holder.....	76
3.2 Experimental Design.....	77
3.2.1 Samples	77
3.2.2 Determination of Sample Size	80

3.2.3 Scanned Samples	88
3.3 Data Analysis.....	90
3.3.1 Spectral Analysis	90
3.3.2 Statistical Analysis.....	92
4. RESULTS AND DISCUSSION	94
4.1 Exploration of Data.....	94
4.1.1 Image Exploration.....	94
4.1.2 Statistical Exploration	96
4.2 Detection of <i>Fusarium</i> spp. in Whole Wheat Kernels	100
4.3 Effects of Canadian Soil Zones.....	106
4.4 Prediction of DON Content from Infection by <i>Fusarium</i> spp.	109
5. CONCLUSIONS.....	118
6. RECOMMENDATIONS	121
7. REFERENCES	124
APPENDIX A: DESIGN OF A GRAIN KERNEL HOLDER.....	136
APPENDIX B: CANADIAN GRAIN COMMISSION’S OFFICIAL GRAIN GRADING GUIDE.....	137
APPENDIX C: SOIL ZONE, LOCATION, AND FUSARIUM DAMAGE PERCENTAGE OF SAMPLES	143
APPENDIX D: SAMPLES USED IN THE EXPERIMENT AS DETERMINED BY A RANDOM NUMBER GENERATOR	146

APPENDIX E: ATR PRESSURE POINT AND SPECTRAL IMAGES WITHIN THE ATR	
PRESSURE POINT	148
APPENDIX F: TYPICAL SPECTRAL IMAGES AND WHEAT KERNEL SURFACES	150

LIST OF TABLES

Table 3.0	Measured kernel sizes.....	77
Table 3.1	Functional group with left and right baseline of FWHM.....	92
Table 4.0	Statistical descriptor results of the integrated values from the exploratory phase.....	98
Table 4.1	Statistical descriptor results of the ratio values from the exploratory phase.....	99
Table 4.2	P-value results of several statistical tests for the integration values.....	102
Table 4.3	P-value results of several statistical tests for the ratio values.....	104
Table 4.4	P-value results of contrast tests performed for ratio values.....	105
Table 4.5	P-value results of soil specific contrast results for integration values.....	107
Table 4.6	P-value results of soil specific contrast results for ratio values.....	108
Table 4.7	Bin sizes for moving average technique.....	110
Table 4.8	Mean, standard deviation, and confidence limits of spectra significantly different than the spectrum of an uninfected kernel.....	114
Table 4.9	Mean values and confidence limits used – elimination of clean pixels using upper or lower confidence limit.....	116

LIST OF FIGURES

Figure 2.0	The electromagnetic spectrum.....	9
Figure 2.1	Potential energy curve as a function of internuclear distance of two vibrational states with corresponding rotational levels.....	16
Figure 2.2	Anharmonic model of the potential energy curve as a function of internuclear distance.....	21
Figure 2.3	Interaction modes of radiation with matter.....	29
Figure 2.4	Attenuated total reflectance.....	31
Figure 2.5	Snell's law.....	31
Figure 2.6	Total internal reflection and snell's law.....	33
Figure 2.7	Evanescent wave properties.....	34
Figure 2.8	Schematic of the components of a FT-IR spectrometer.....	36
Figure 2.9	Michelson interferometer.....	39
Figure 2.10	Phase difference between the fixed (solid line) and moveable (dashed line) mirrors and resultant output from the interferogram.....	40
Figure 2.11	Interferogram and corresponding spectrum.....	43

Figure 2.12	Fusarium damaged kernels (FDK) artificially infected at different life stages.....	59
Figure 3.0	Common FT-IR accessories and corresponding sample types.....	75
Figure 3.1	The distribution of <i>F. Graminearum</i> across the prairie provinces.....	79
Figure 3.2	Histogram of protein amide one integrations.....	82
Figure 3.3	SAS power procedure results for protein amide one.....	84
Figure 3.4	SAS power procedure results for lipid carbonyl.....	85
Figure 3.5	G*Power 3 result for the protein amide one functional group.....	86
Figure 3.6	G*Power 3 result for the lipid carbonyl functional group.....	87
Figure 3.7	Top, side, and bottom view of wheat kernel body.....	89
Figure 4.0	Spectrum of an infected kernel that is significantly different than the typical spectra of an uninfected kernel.....	112

1. INTRODUCTION

1.1 Canadian Grain Industry: Present and Future

Canada is one of the leading grain producing countries in the world, with a total production of 106.2 million tonnes of field crops (including wheat, oats, barley, rye, flaxseed, canola, corn, peas, and soybeans) in the 2009-2010 crop year (Statistics Canada, 2010). Grain is produced on farm, harvested, and trucked to small country elevators. The staff members at the elevators evaluate fundamental quality parameters, such as protein and moisture content, in order to sort, clean, and dry the grain, depending on its condition when received. The grain from the elevators is then transported, via railcars, to large terminal elevators.

The terminal elevators load grain onto ships and vessels for export. On average, Canada exports 21.4 million tonnes of grain annually to numerous countries around the world. The majority of the exported grain is wheat, Canadian Western Red Spring (CWRS) wheat being the prevalent crop. Prior to export, trained inspectors from the Canadian Grain Commission (CGC) visually grade the grain in order to ensure quality control so customers are receiving their desired end product (Canadian Grain Commission, 2009).

The grain industry is significant to Canada's economy, employing two million people across the country and contributing approximately 8.2% of Canada's gross domestic product annually (Agriculture and Agri-Food Canada, 2011). Grading grain is based on determining fundamental quality parameters to ensure a reputable end product;

the more desirable the quality parameters within the grain batch are, the higher the grain is graded. This grading process is the basis of the Canadian grain industry and sets Canada apart from other competitors in the business. In order to compete in the ever growing market, and maintain a distinguished reputation, Canada must continue to ensure grain quality while increasing production (Canadian Wheat Board, 2008).

Currently, the Canadian Wheat Board (CWB) develops an annual marketing strategy and sales plan for all of the wheat and barley commodities exported from Canada. The CWB does this by comparing “the world supply, demand, and price forecasts” of different countries in an attempt to maximize profits. In order for a customer to buy a Canadian grown grain product, it must be purchased through the CWB (Canadian Wheat Board, 2011). The price the grain is sold for reflects the grade of the grain; a higher grade corresponds to a higher price.

Despite the fact that Canada’s grading system is recognized as one of the best worldwide (Canadian Wheat Board, 2008), grain is graded primarily on visual methods. These include, but are not limited to, kernel colour and shape, varietal purity, soundness, and vitreousness (Canadian Grain Commission, 2011). These visual methods make it difficult to achieve consistent grades among different inspectors as the grading process is subjective and tedious. In order to maintain Canada’s reputable status of providing quality grain, and achieve a faster throughput of grain commodities, consistent, rapid, and precise techniques must be developed to objectively determine the quality parameters of grain for grading purposes.

1.2 Fusarium Fungi: The Need for an Objective Measurement

Numerous fungi can affect the quality of grain, downgrading it and potentially rendering it unfit for consumption. One of the most common fungi that infect grain in Canada is different species of *Fusarium*. Infection of grain by *Fusarium* spp. is commonly associated with high moisture conditions during crop growth (Clear and Patrick, 2010). This fungus has the potential to destroy a healthy crop in a matter of weeks (Goswami and Corby, 2004), and is responsible for reducing crop yields as much as 50 percent in some years (Clear and Patrick, 2010). *Fusarium* infected grain leads to undesirable end products such as “poor flour colour, weak dough properties, and unsatisfactory bread quality” (Dexter, et al., 1996). Additionally, *Fusarium* infected grains may contain a family of mycotoxins known as trichothecenes, the most common toxin of this family being deoxynivalenol (DON) (Shahin and Symons, 2011). If DON is ingested, it can cause vomiting, diarrhoea, and weight loss in humans and animals (Abramovic et al., 2007). The health, safety, and economic risks associated with this fungus are significant; therefore, the need to accurately determine its existence is essential for Canada to maintain a strong presence in the grain industry.

There are four strains of *Fusarium* in North America: *F. graminearum*, *F. culmorum*, *F. avenaceum*, and *F. crookwellense*. In Canada, and much of the world, *F. graminearum* is the most predominant species. This is due to its high genetic variability, tolerance to different temperature and moisture conditions, aggressive nature (Osborne and Stein, 2007), and its ability to produce several different mycotoxins, such as deoxynivalenol (Clear and Patrick, 2010).

In order to determine whether a batch of grain contains any *Fusarium* spp., CGC inspectors, or qualified inspectors as designated by the CGC, remove a sample from the batch and visually inspect it for kernels that appear to be infected. If the inspectors are unsure about any of the kernels, a 10× microscope is used to confirm or deny the presence of *Fusarium* spp. The percentage of *Fusarium* damaged kernels in the sample is taken to be representative of the entire batch and a grade of the grain is assigned based on this, as well as several other aforementioned quality parameters. Depending on the extent of kernels damaged by the fungus, the batch may be downgraded to feed or discarded (Canadian Grain Commission, 2011).

The current methods used to detect *Fusarium* spp. are archaic and subject to error. Factors such as fatigue, working conditions, level of knowledge, etc., may result in inaccurate classification of kernels that are infected with *Fusarium* spp. This methodology is also tedious; consuming an inspector's time and consequently driving costs up for the company employing the individual. Furthermore, the sample that is withdrawn may not actually be representative of the entire batch, resulting in misclassification of grain grade. In order to relieve the grain industry of these challenges, it is desirable to provide an objective measurement to classify kernels infected with *Fusarium* spp. in an accurate and rapid manner.

1.3 Fourier Transform Infrared Microspectroscopy as a Viable Technology for Grain Quality Assessment

Fourier Transform Infrared (FT-IR) microspectroscopy has been extensively tested in grain industry applications to validate its technical capabilities in relation to determining quality parameters of grain commodities. To date, these validation studies have been

successful in characterizing the chemical parameters of the grain commodities (Section 2.6.2). The reason for the success of FT-IR microspectroscopy and why it continues to be a method of choice in agricultural studies arise from several advantages the technology possesses:

1. Results obtained from FT-IR microspectroscopy are sensitive and reliable. The sensitivity of the method arises from the Fast Fourier Transform (FFT) algorithm used, as well as the capability of the instrumentation to allow radiation to make contact with the samples. Reliability comes from the fact that the FT-IR microspectroscopy process is highly automated with little user-machine interaction (Griffiths and de Haseth, 2007).
2. FT-IR microspectroscopy is a rapid measurement technique. The reason for this comes from the instrument design employed in FT-IR microspectroscopy that allows spectral information to be gathered from numerous wavelength ranges simultaneously (Griffiths and de Haseth, 2007).
3. FT-IR microspectroscopy provides a non-destructive analysis of a variety of samples. Samples do not need to be pre-treated prior to scanning, which further increases the speed of the technique. Additionally, samples in a number of different states (solid, powder, liquid) can be analyzed by this method (Singh et al., 2011).
4. The instrumentation employed by FT-IR microspectroscopy is relatively simple, with only one moving component (Griffiths and de Haseth, 2007). The risk of mechanical degradation of the instrumentation is relatively low due to the minor

movement the machinery undergoes when in use. Therefore, cost of maintenance of the instrumentation is minimal to none.

5. FT-IR microspectroscopy is an optical method and thus does not release any harmful emissions upon operation. Due to this fact, the operator of the FT-IR microspectroscopic instrument does not need to wear any personal protective equipment while the machine is in use. Moreover, operation of the instrument requires little training and the entire process of collecting a sample can be performed by one user.

Despite the widespread use and technical advantages of FT-IR microspectroscopy that make it a viable technique for grain quality assessment, more research could be done into its applications with grain commodities to increase its utility in the agricultural industry.

1.4 Objectives of Research

In the past decade, there have been numerous research studies published that explore the use of FT-IR microspectroscopy as a detection technique for the presence of *Fusarium* spp. in a variety of grain kernels in order to overcome the limiting factors (Section 1.2) involved with the current grain grading processes. These experiments generally use an accessory, such as attenuated total reflectance (ATR) or diffuse reflectance (DR), in conjunction with FT-IR microspectroscopy in order to provide a sensitive detection mode. However, all of the studies to date have pre-treated the kernels via grinding or sectioning to ensure that contact between the sample and the ATR crystal could be made. This introduces a tedious step in the data collection process.

The focus of this thesis is determination and evaluation of the feasibility of FT-IR microspectroscopy, coupled with ATR, to rapidly detect and measure *Fusarium* spp. and DON content, respectively, on whole wheat kernels. FT-IR microspectroscopy was chosen as it provides a rapid, sensitive, non-destructive, objective measurement of the chemical characteristics present in the sample. It is hoped that gaining a better understanding of the applicability of FT-IR microspectroscopy with whole wheat kernels, as it relates to *Fusarium* infection and DON content, can help meet the grading demands of the agricultural industry in the future.

The specific objectives of this research are:

1. to determine if *Fusarium* spp. can be detected in whole wheat kernels via FT-IR microspectroscopy coupled with an ATR accessory by identifying the spectral parameters that indicate the presence of the fungus;
2. to recognize and classify any differences in the ability to identify *Fusarium* spp. in whole wheat kernels among the Canadian soil types;
3. to develop an equation that correlates the amount of DON content (ppm) to one or more of the fundamental spectral parameters that indicate the presence of *Fusarium* spp.; and
4. to carry out an investigation to test the validity of the developed equation to determine if it can be used in future microspectroscopic experiments to relate the identified spectral parameters to DON content.

The remainder of the thesis is organized as follows: a literature review, materials and methods, results and discussion, conclusion, and recommendations. The literature

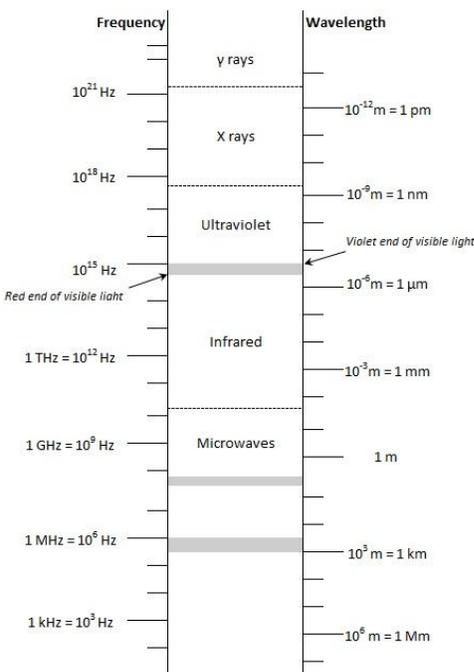
review section will cover the theory behind FT-IR spectrometers and an overview of the current applications of FT-IR microspectroscopy, as it relates to the detection of *Fusarium* spp. and DON content. The materials and methods section describes the instrumentation used and sample design of the experiment. The materials and method section also outlines the method used to analyze the data. The results and discussion section presents the results of the obtained data as they relate to the objectives of the thesis. A corresponding analysis of the results is also given. The conclusion section summarizes the accomplishments of the thesis. Finally, the recommendations section gives direction for future research in the area of FT-IR microspectroscopy with ATR for the detection of *Fusarium* spp. and DON content prediction in whole wheat kernels.

2. LITERATURE REVIEW

2.1 History and Background of Fourier Transform Infrared Spectroscopy:

The infrared portion of the electromagnetic spectrum extends from 780 to 100,000 nm and is comprised of three sub-regions: near-infrared (780 to 2,500 nm), mid-infrared (2,500 to 25,000 nm) and far-infrared (25,000 to 100,000 nm). Fig. 2.0 depicts the infrared region in relation to the rest of the electromagnetic spectrum. This spectral region was discovered in the year 1800 by the astronomer, Sir William Herschel, while attempting to find the cause of heat formation in his telescope (Bosco, 2010).

Figure 2.0 The electromagnetic spectrum. *Source:* Laidler et al., 2003



Spectroscopy can be defined as, “the study of the interaction between electromagnetic radiation and matter” (Laidler et al., 2003). The foundation of spectroscopy occurred in 1665 when Sir Issac Newton demonstrated that a spectrum of

colours, ranging from red to violet, could be obtained by passing white light through a prism (Laidler et al., 2003). The first spectrometer was developed in 1850 by Bunsen and Kirchoff, and was able to operate solely in the visible range. By 1900, W.W. Colbenz acquired the first IR spectrum (Murray, 2003), and was able to demonstrate that similar compounds exhibit similar characteristic spectra (Bosco, 2010).

Adopting Laidler et al's (2003) definition of spectroscopy, IR spectroscopy can be defined as the study of matter by subjecting a sample to IR radiation and observing the interaction effects of the radiation and the sample in question. In 1936 the first IR spectrometer was put into industrial settings (Barnes and Gore, 1949). Advances in statistics and algorithms, computer technology, and instrumental components throughout the years have made IR spectroscopy a practical and widely applied measurement technique.

Fourier Transform Infrared (FT-IR) spectroscopy is a type of IR spectroscopy that utilizes unique instrumentation and applies a Fast Fourier Transform (FFT) algorithm to the spectral results. The spectrum obtained from FT-IR spectroscopy is essentially a molecular fingerprint of the sample in question, reflecting its molecular background and environment. FT-IR spectroscopy was being applied to microorganisms as early as the 1950s; however, only within the last few decades has FT-IR spectroscopy been realized as a powerful characterization tool for complex biological systems (Duygu et al., 2009).

There are several advantages involved with using FT-IR spectroscopy, which have led to significant replacements of conventional techniques (such as diffraction grating) in spectroscopy (Singh et al., 2011). These include, but are not limited to, fast

scanning speed and greater accuracy of results. These advantages arise from the instrumentation, namely the interferometer, employed in the method as well as the ability of the FFT algorithm to characterize samples.

The instrumentation employed by FT-IR spectroscopy allows for a rapid scan of the sample in question. This is primarily due to the multiplex and throughput advantages that FT-IR spectroscopy has over conventional techniques. The multiplex advantage arises from the fact that spectral information can be obtained simultaneously from all wavelengths. This allows for overall higher signal-to-noise (SNR) ratios, which increases the sensitivity and speed of the FT-IR scan, when compared to traditional spectroscopic methods (Griffiths and de Haseth, 2007). The throughput advantage results from the fact that the beam width of FT-IR instruments does not need to be limited. This allows a higher intensity of radiation to be passed through the sample and, consequently, a faster throughput time (Singh et al., 2011).

FT-IR spectroscopy is often described as a non-destructive technique, as the samples employed do not need to be pre-treated prior to scanning. This lack of sample preparation further increases the speed of the system and reduces error due to the elimination of “weighing and dilution operations” (Singh et al., 2011). A wide variety of samples can be scanned using this method, as spectra can be obtained regardless of the state (solid, powder, liquid) of the substance in question. Additionally, the spectrometer can be adjusted to allow scanning of extremely small samples (Duygu et al., 2009).

FT-IR spectroscopy utilizes the FFT algorithm when obtaining spectra, thus making it a faster technique. It has the ability to characterize different macromolecules

including lipids, carbohydrates, and DNA, as well as many others. This technique is so specific, that a scan can be done on an entire sample and differentiation of the different components of the sample can be made (Duygu et al., 2009).

Due to these advantages, FT-IR spectroscopy has been successfully employed to detect a wide range of biological characteristics in a vast number of different industries. These include, but are not limited to, the fields of agriculture, pharmacy, microbiology, medicine, and food processing (Santos et al., 2010). In order to understand and appreciate this powerful analytical technique, the theory which lies behind its operating principle must be understood.

2.2 Theory:

FT-IR spectroscopy is based on the theory of molecular vibrations and overtones in which bands at specific frequencies appear in a spectrum due to characteristic vibrational modes of different molecular groups. Consequently, spectra obtained from FT-IR spectroscopy provide a static image of the organic constituents of the sample in question. Organic matter is composed of atoms; mostly carbon, hydrogen, oxygen, and nitrogen, which covalently bond with one another to form molecules. These covalent bonds have unique characteristics such as length, strength, and direction which are dependent upon the atoms that form the bonds, as well as the type of bond formed. In addition to their physical characteristics, the interaction between the atoms and the bonds allow the bonds to vibrate at a certain fundamental frequency and permit the entire molecule to rotate and move (Laidler, et al., 2003).

The electromagnetic spectrum, and consequently IR radiation, is composed of a series of photons which travel at the speed of light ($c = 3.0 \times 10^8$ m/s). These photons can be described in terms of frequency (ν) and wavelength (λ) by the expression:

$$c = \lambda\nu \quad (2.1)$$

Planck's constant ($h = 6.63 \times 10^{-34}$ Js) enables the energy of a photon (ΔE) to be related to its frequency and wavelength by the expression:

$$\Delta E = h\nu = \frac{hc}{\lambda} \quad (2.2)$$

Photons will only be absorbed by molecules if the change in energy between the energy levels of the molecule identically matches the energy of the photon (Carter, 1998). The absorption intensity is dependent upon the change in the dipole moment of the molecule, which is a result of its vibration (Åmand and Tullin, 1999). The IR spectrum displays the energies of these vibrations.

2.2.1 Vibrational-Rotational Energy of Molecules:

Due to their simplicity, diatomic molecules can be used to illustrate molecular vibrational and rotational energy. A vibrating diatomic molecule consists of bond stretching, compressing, and restoration to the equilibrium position. Assuming this bond motion occurs in a classical mechanical way, the restoring force, F , acting upon the molecule can be modeled by Hooke's law:

$$F = -k\Delta r \quad (2.3)$$

Where k is the force constant (N/m) that acts to bring the atoms to their equilibrium position (r_e) and Δr is the distance the atoms have moved from their equilibrium position. The two atoms of the diatomic molecule will travel away from each other, stretching the bond, until they reach a certain point upon which the restoring force causes them to reverse their motion and travel towards each other. The atoms will then continue to travel towards each other until their repulsive forces overcome the inward motion and drive them to move apart from each other again. This oscillating motion is known as harmonic oscillation and constitutes one cycle of a vibrational mode (Carter, 1998). Due to the oscillating nature of atomic vibrations, Hooke's law can be used to model the frequency of the vibrational mode:

$$\nu = \frac{1}{2\pi} \sqrt{\frac{k}{\mu}} \quad (2.4)$$

Where μ is the reduced mass (kg):

$$\mu = \frac{m_1 m_2}{m_1 + m_2} \quad (2.5)$$

m_1 and m_2 being the two masses (kg) in a diatomic molecule (Laidler et al., 2003).

Due to the fact that energy cannot continuously be absorbed by matter, Max Planck proposed a quantum-mechanical approach in which oscillations are characterized by equally spaced vibrational energy levels (Osborne et al., 1993). Based on this model, the potential energy of a diatomic harmonic oscillator, for different vibrational quantum numbers, can be found via the equation:

$$\Delta E_V = h\nu_0 \left(v + \frac{1}{2} \right) \quad (2.6)$$

where ν_0 is the natural frequency of motion (Hz), and v is the vibrational quantum number ($v = 0, 1, 2, \dots$) (Laidler et al., 2003). Eq. 2.6 is similar to Eq. 2.2 for the energy of a photon; however, Eq. 2.6 shows that atoms are never completely at rest. A zero-point energy occurs for each bond. This is the lowest possible energy of a bond and can be found by setting $v = 0$ and ultimately obtaining the equation, $\Delta E_v = \frac{1}{2} h\nu_0$ (Laidler et al., 2003).

The rotational component of a molecule's spectrum can also be modeled using a diatomic molecule. Rotation of a molecule can only be achieved when the molecule is in a gaseous phase (ie. water vapour). The model characterizes the diatomic molecule as a rigid rotor, in which the bond is a rod of negligible weight. The energy component of such a model is given by:

$$E_J = BJ(J + 1)hc \quad (2.7)$$

where J is the rotational quantum number ($J = 0, 1, 2, \dots$), and B is the rotational constant (no unit):

$$B = \frac{h}{8\pi^2 Ic} \quad (2.8)$$

The moment of inertia, I (kg m^2), is defined as:

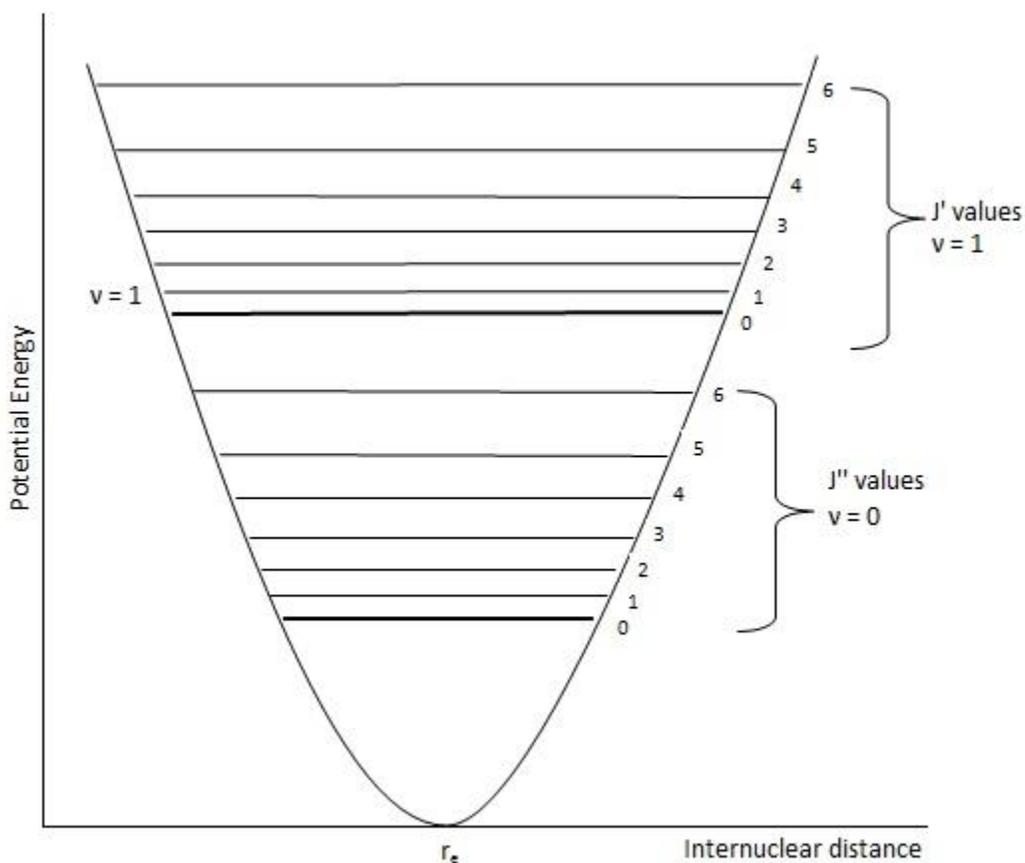
$$I = \mu r^2 \quad (2.9)$$

where μ is the reduced mass (Eq. 2.5) and r is the bond length (m). Thus, the vibrational (Eq. 2.6) and rotational (Eq. 2.7) energies of a diatomic molecule give a combined potential energy of:

$$E_{v,J} = hv_0\left(V + \frac{1}{2}\right) + BJ(J + 1)hc \quad (2.10)$$

Eq. 2.10 operates based on the assumption that vibrational and rotational energies are separable. Hence, if $\Delta v=0$, the molecule will exhibit pure rotational motion. If $\Delta v \neq 0$ the molecule has both vibrational and rotational motions. Fig. 2.1 models this superposition for a harmonic, diatomic molecule with two vibrational energy states ($v = 0, 1$) and six rotational energy levels for each state ($J = 1, 2, 3, 4, 5, 6$).

Figure 2.1 Potential energy curve as a function of internuclear distance of two vibrational states with corresponding rotational levels. *Source:* Laidler et al., 2003



2.2.2 Molecular Degrees of Freedom and Spectroscopic Selection Rules:

A molecule with N atoms possesses $3N$ degrees of freedom, which describe its translational, rotational, and vibrational movement. Of these degrees of freedom, three will always be translational and these correspond to the three axes of the Cartesian coordinate system. The number of rotational degrees of freedom depends on whether the molecule is linear or non-linear, and the number of vibrational modes is equal to $3N$ minus the number of translational and rotational modes (Carter, 1998).

In a nonlinear molecule, three degrees of freedom are rotational. This is due to the fact that three moments of inertia exist about the three axes of the Cartesian coordinate system. In linear molecules, the axis in which the molecule lies possesses no moment of inertia, and thus only two degrees of freedom are rotational. The vibrational degrees of freedom (dof) of linear and nonlinear molecules can be given by the following equations, respectively: $\text{dof} = 3N - 5$ and $\text{dof} = 3N - 6$. These equations account for the translational and rotational movement of the molecules (Carter, 1998). As an example, benzene (C_6H_6) is a nonlinear molecule consisting of twelve atoms. Using the equation $\text{dof} = 3N - 6$, one can see that benzene possesses thirty vibrational degrees of freedom [$\{(3 \times 12) - 6\} = 30$].

In general, the number of vibrational modes corresponds to the number of bands seen in the IR spectrum. However, this is not always the case and there are several spectroscopic selection rules that govern whether the vibrations induced in a molecule will be displayed in an IR spectrum. These selection rules are based upon the symmetry

of the molecule, which ultimately determines the type and direction of rotations and vibrations that can occur. Only a few of the selection rules will be discussed. Additional details regarding the complete set of spectroscopic selection rules can be found in other texts.

The first spectroscopic rule of concern states: in order for vibrations to be produced by IR radiation, there must be a change in the permanent dipole moment of the molecule (Carter, 1998). The intensity of the IR spectrum is proportional to the change in the permanent dipole moment of the molecule; hence, the greater the change in the permanent dipole moment of a molecule, the higher the intensity of the IR spectrum. The permanent dipole moment of a molecule can be defined as the effective charge at the positive and negative ends of the molecule, multiplied by the distance between them. Therefore, any diatomic molecule that does not possess a dipole moment (e.g. O₂, N₂, etc.) is not capable of displaying infrared spectra (Laidler et al., 2003).

A change in the dipole moment for a diatomic molecule is relatively simple, as there are only two atoms that produce a single vibrational frequency. However, situations in which there are more than two atoms in a molecule are subject to restrictions. The dipole moment of these molecules is a vector that can be broken into x, y, and z components, corresponding to the Cartesian coordinates. If the dipole moment of the x component is of a different symmetry than that of the normal vibrational mode, there will be no change in the x dipole moment. Hence, if all of the x, y, and z components have different symmetries than those of the corresponding normal vibrational modes, there will be no change in the dipole moment of the molecule and there will be no vibrational-rotational spectrum. This is the second selection rule of consideration, known as the

symmetry restriction, which states: “there will be a vibrational-rotational spectrum corresponding to a normal mode of vibration only if that mode belongs to the same symmetry species as one or more of three coordinates x , y , and z ” (Laidler et al., 2003).

There are cases in which the number of bands in the IR spectrum is not equal to the calculated vibrational degrees of freedom. This may be attributed to the aforementioned lack of change in the permanent dipole moment of the molecule or the discrepancy in symmetry between the vectors of the dipole moment and the normal vibrational mode. Other reasons for this discrepancy include silent modes (Carter, 1998), weak absorption, poor instrument resolution, and absorption occurring outside of the specified wavelength range (Bradner et al., 1966).

The final spectroscopic selection rule of consideration states: energy can only transition from one fundamental state to the next, hence $\Delta V = \Delta J = \pm 1$. Under strict harmonic motion, with both vibrational and rotational energy involved, a change in the vibrational state can only occur to the next level if there is a corresponding change in the rotational level. However, transitions that digress from this rule occasionally do occur. These deviations can be explained using an anharmonic model to describe the behaviour of diatomic molecules (Laidler et al., 2003).

2.2.3 Anharmonic Model:

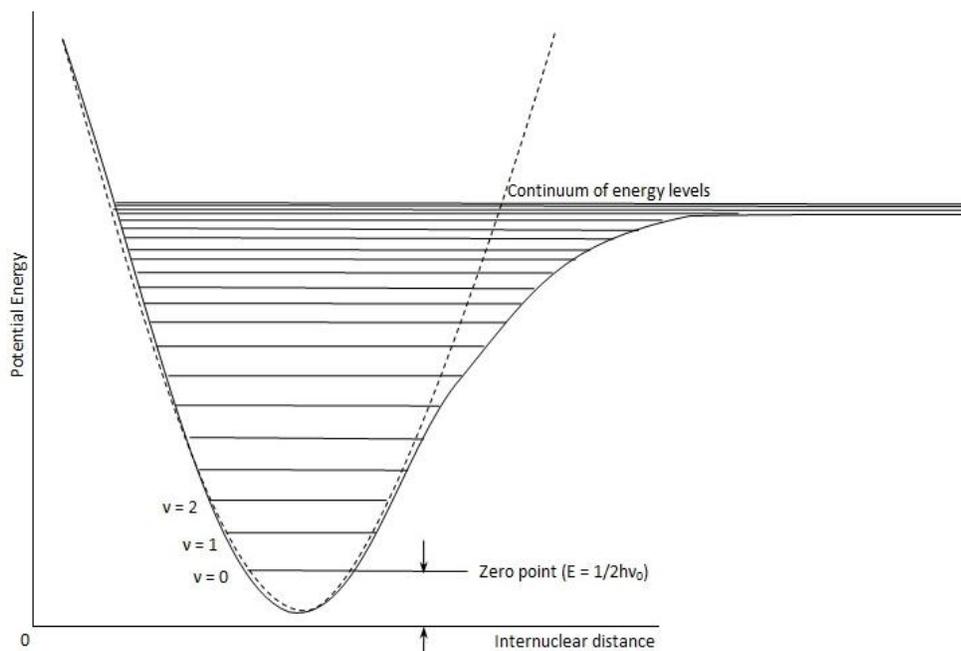
The assumption of harmonic oscillation is useful as an approximation model for molecular vibration and rotation. However, this is simply an approximation and cannot be used to model all vibrational and rotational states. A representation known as the

anharmonic model takes into account several features that are exhibited by true diatomic molecules, which the harmonic model fails to take into consideration.

Fig. 2.2 is a diagram of the potential energy of the harmonic model (dashed line) and anharmonic model (solid curve), as a function of internuclear distance. This plot shows the important differences between the two models. It can be seen from Fig. 2.2 that, at lower energy levels, the harmonic oscillation model is able to accurately predict the vibrational and rotational energy levels of molecules. However, at high energy states and larger internuclear separations, the discrepancy between the harmonic and anharmonic models is vast.

The harmonic oscillator model gives a true parabolic shape when potential energy is plotted as a function of internuclear distance. This shape implies that potential energy continually increases with increasing internuclear separation, which signifies that molecular bonds cannot be broken. However, in reality, molecular bonds can and will break when the internuclear distance between the atoms increases to a critical point. This is accounted for in the anharmonic model and can be seen as the reduction in the difference between energy levels as the vibrational quantum number, and consequently the potential energy, increases. The differences eventually become so small that a maximum potential energy is reached, which signifies that the molecule has dissociated into atoms (Laidler et al., 2003).

Figure 2.2 Anharmonic model of the potential energy curve as a function of internuclear distance. *Source:* Laidler et al., 2003



The potential energy of anharmonic diatomic molecules cannot be modeled using Hooke's law. However, this energy can be closely predicted by the Morse potential function:

$$E_p = D_e(1 - e^{-ax})^2 \quad (2.11)$$

where x is the extension of the bond from its equilibrium position (m), D_e is the dissociation energy (J) and a is a constant. The energy of the vibrational levels must also be modified to fit the anharmonic model; this results in the following equation:

$$\Delta E_v = hv_0 \left(v + \frac{1}{2} \right) - hv_0 \left(v + \frac{1}{2} \right)^2 w + hv_0 \left(v + \frac{1}{2} \right)^3 w' - \dots \quad (2.12)$$

where w and w' are anharmonicity constants.

The Morse potential function accounts for the closely spaced potential energy levels at large internuclear distances of the anharmonic model. This continuum of energy levels yields vibrational quantum numbers of less than the whole integer values they were defined as in Eq. 2.6. The less than unity vibrational quantum numbers can explain transitions digressing from the selection rules. These deviations are commonly seen as overtones and combination bands in the IR spectrum (Bokobza, 1998).

2.2.4 Fourier Transforms

The Fourier transform is a mathematical method that transforms a signal, which is originally in the time domain, into a representation of the signal in the frequency domain. This method is widely used as a means of simplifying and understanding complex processes. The Fourier transform algorithm has evolved throughout time to provide faster computation speeds in order to solve convoluted problems (Dahamel and Vetterli, 1999). Three different Fourier algorithms are discussed in the following section: the classical Fourier transform, the discrete Fourier transform, and the fast Fourier transform.

The classical Fourier transform operates by multiplying every interval point of a cosine- wave interferogram by a corresponding point of an analyzing cosine wave of unity value (Griffiths and de Haset, 2007). The discrete Fourier transform applies a finite duration to the data obtained from the classical Fourier transform. The fast Fourier transform is a combination of the classical and discrete Fourier methods that significantly reduces the computation time required to process results. The fast Fourier transform is the algorithm that is currently used in spectroscopic applications. This method has influenced

the fields of science and engineering by providing the simplest means to compute spectral data (Reddy, 1998).

Classical Fourier Transform: the classical Fourier transform is defined by:

$$H(\Omega) = \int_{-\infty}^{\infty} h(t)e^{-j\Omega t} dt, \quad j = \sqrt{-1} \quad (2.13)$$

where $h(t)$ is the continuous-time signal, and Ω is the frequency in radians per second ($\Omega = 2\pi f$). If the signal described in Eq 2.13 is sampled at set time scale intervals in order to guarantee signal recovery, the Fourier transform is given by:

$$H(\omega) = \sum_{n=-\infty}^{\infty} h(nT)e^{-j\omega n} \quad (2.14)$$

where ω is the normalized frequency in radians ($\omega = \Omega T$), T is the time in seconds, n is the number of times the signal has been sampled, which is an infinite value in this case. It can be seen from Eq 2.14 that $H(\omega)$ is a periodic function with a period of 2π (Reddy, 1998).

Optical retardation (Section 2.4.2) can play a role in the resultant value of the classical Fourier transform, depending on the frequency relation of the interferogram and the analyzing wave. If the frequency of the analyzing wave is significantly different than the frequency of the interferogram, optical retardation is not an influencing variable. However, if the frequency of the analyzing wave is close to that of the interferogram, optical retardation becomes a factor. As optical retardation increases, the resulting frequency oscillation of the classical Fourier transform product becomes higher, and consequently, the spectral resolution becomes higher as well (Griffiths and de Haseth, 2007).

In order to compute a complete spectrum over an infinite range, the sampling frequency must be greater than or equal to twice the bandpass of the system. This is known as the Nyquist criterion. The computation time involved with the classical Fourier transform and the Nyquist frequency is quite tedious, even for limited spectral ranges. Recursive relationships can assist with reducing computation time, the most prevalent one being the Chebyshev formula (Griffiths and de Haseth, 2007). The Chebyshev differential equation is: $(1-x^2)y'' - xy' + m^2y = 0$ and can be used to generate polynomials in order to provide accurate approximations (Biswal, 2008). Despite this assistance, it can be proven that each wavenumber requires a total of N^2 operations in order to produce the desired output, where N is the total number of times the signal has been sampled (infinite in this case) (Duhamel and Vetterli, 1999). Even with advancing computer technology and speed, the classical Fourier transform method is not practical for spectroscopic applications and the discrete Fourier transform algorithm was derived (Griffiths and de Haseth, 2007).

Discrete Fourier Transform: the discrete Fourier Transform (DFT) is an algorithm that is used to truncate the classical Fourier transform in order to estimate the signal at uniformly spaced intervals. A DFT pair can be derived from the classical Fourier transform by limiting the sequence to a finite duration of N samples. This pair is a combination of a synthesis equation (Eq. 2.15) and an analysis equation (Eq. 2.16), which is similar to the measured and analyzing waves of the classical Fourier transform, respectively:

$$h(nT) = \frac{1}{N} \sum_{k=0}^{N-1} H\left(\frac{k}{NT}\right) e^{\frac{j2\pi nk}{N}}, \quad n = 0, 1, \dots, N - 1 \quad (2.15)$$

$$H(k/NT) = \sum_{n=0}^{N-1} h(nT)e^{\frac{-j2\pi nk}{N}}, \quad k = 0, 1, \dots, N - 1 \quad (2.16)$$

where N is a finite number of samples and $1/N$ is a normalizing factor (Reddy, 1998).

The DFT simplifies the overall process involved with the classical Fourier transform, due to its finite domain. However, truncating the data does result in poorer estimates of the measured data. Ultimately, the DFT algorithm still requires N^2 multiplications for each wavenumber in order to produce an output spectrum (Griffiths and de Haseth, 2007). Computation time is slightly accelerated by specifying the finite domain, yet another algorithm is required in order to make the Fourier method practical in spectroscopic settings.

Fast Fourier Transform: the fast Fourier transform (FFT), first described by Cooley and Tukey in 1965 (Dahamel and Vetterli, 1999), is capable of reducing the number of computations involved with the classical and discrete Fourier transform methods. One can define a parameter, W_N , to be: $W_N = \exp\left(\frac{-j2\pi}{N}\right)$ such that the following two properties can be derived:

$$W_N^{k(N-n)} = (W_N^{kn}) \quad \text{Property 1}$$

$$W_N^{kn} = W_N^{k(n+N)} = W_N^{(k+N)n} \quad \text{Property 2}$$

The first property reduces the number of computations of the classical and discrete Fourier transform methods by a factor of two and the second property can be used to reduce the number of operations even further. Using this parameter, the FFT can be described by the equation:

$$H(k) = \sum_{r=0}^{\left(\frac{N}{2}\right)-1} h(2r)(W_N^2)^{rk} + W_N^k \sum_{r=0}^{\left(\frac{N}{2}\right)-1} h(2r+1)(W_N^2)^{rk} \quad (2.17)$$

where r is the number of discrete wavenumbers (Reddy, 1998). The above equation can be expressed in spectroscopic terms by the equation (Griffiths and de Haseth, 2007):

$$B(\tilde{\nu}) = \sum_k^{N-1} S_0(k)W^{rk} \quad (2.18)$$

The FFT algorithm is often referred to as the “divide and conquer” method. When computing problems of this nature, the original problem is broken down into several sub-problems such that the number of operations and complexity of the sub-problems is less than that of the original problem. This condition is often accomplished by using recursive formulas in the sub-problems which reduces the complication associated with the original problem (Dahamel and Vetterli, 1999). Eq. 2.18 achieves this by its ability to be expressed in a general matrix form (Griffiths and de Haseth, 2007).

The FFT algorithm cuts down the number of operations of the classical and discrete Fourier methods from N^2 operations to $N \log_2 N$ (Dahamel and Vetterli, 1999). The FFT technique enables low resolution spectra to be produced in a matter of seconds, and is therefore practical for high resolution spectral applications. The only criterion limiting the application of the fast Fourier transform in spectroscopic applications is that the number of data points in the interferogram must be an integer power of two (i.e. $N = 2^n$) (Griffiths and de Haseth, 2007). This is easily achieved by the traditional focal plane arrays and user interface setting on FT-IR instruments and software. Additional details regarding the vibrational-rotational theory of FT-IR spectroscopy, as well as the mathematical background behind producing an interferogram may be found in Carter,

1998; Laidler et al., 2003; Griffiths and de Haseth, 2007; as well as many other spectroscopic and chemistry texts.

2.3 Measurement Modes:

There are several different effects that can occur when infrared light encounters a biological sample, depending upon the nature of the sample. These can be broadly categorized as absorbance, transmittance, and reflectance. In general, transparent materials will absorb and transmit radiation and solid, opaque materials will reflect radiation.

2.3.1 Interaction with Matter:

Absorption of light by a sample involves removal of energy from the incident beam. The amount of energy removed from the radiant beam depends on the physical properties of the sample being struck. The absorption (A) of light by a material is quantitatively described by the Lambert-Beer law:

$$A = \log_{10} \frac{I_0}{I} = \epsilon cl \quad (2.19)$$

where I_0 is the intensity of light (J) prior to passing through the substance, I is the intensity of light (J) after passing through the substance, ϵ is the absorption coefficient of the substance ($\text{dm}^3/\text{mol}\cdot\text{cm}$), c is the concentration of the substance (M), and l is the length of the light path (m) (Laidler et al., 2003).

Transmittance (T) refers to the amount of incident light that passes through a sample, or the amount of incident light that is not absorbed by the sample. The value

percentage transmittance ($T\%$) is more commonly used and can be described by the equation:

$$T\% = 100T = \frac{100I}{I_0} \quad (2.20)$$

as can be seen from Eq. 2.20, transmittance is a ratio of the intensity of light after it has passed through a sample to the intensity of light prior to passing through a sample.

Substances employed using the transmittance measurement mode should have minimum light scattering characteristics in order to minimize the amount of reflected light.

Transmittance and light absorption can be related to one another with the following equation:

$$A = \log_{10} \frac{1}{T} \quad (2.21)$$

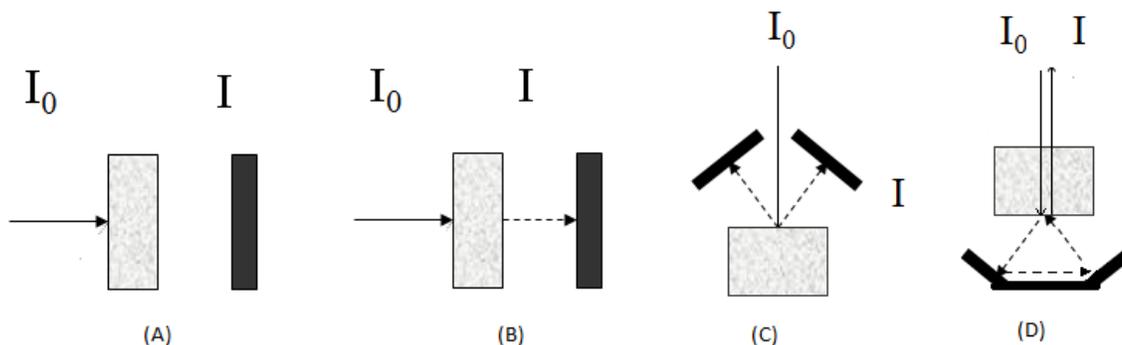
Reflectance occurs when a beam of light encounters a sample and is projected off of its incident surface. Depending upon the type of sample used, diffuse reflectance (DR) may occur. This is when the nature of the sample causes a reflection to appear as though it is coming from multiple angles of the incident surface (Murray, 2003). Reflective measurements are most commonly used on solid, opaque samples. Reflectance is inversely proportional to absorbance and transmittance, hence as the amount of surface reflection increases, the amount of absorption, and consequently transmittance decreases.

Another measurement mode, known as transflectance, exists. As its name implies, transflectance involves a combination of transmittance and reflectance modes.

Transflectance allows the incident light to transmit through the sample twice. This is achieved by placing a reflective material, such as gold or steel, on the side of the sample

which is opposite the incident light source (Singh et al., 2011). An illustration of the different interaction modes of light with matter is shown in Fig. 2.3.

Figure 2.3 Interaction modes of radiation with matter (a) absorption; (b) transmittance; (c) reflection (d) transflection. *Source: Murray, 2003*



Spectra of measured substances are plotted as a function of intensity versus frequency. Percentage transmittance and absorption are generally employed as a measure of the intensity, and wavenumber (the number of wavelengths per unit length) as the frequency. The wavenumber gives information about the type of functional groups present in the molecule, and the intensity gives an indication of the quantity of the present functional groups (Bradner et al., 1966).

2.3.1 Focal Plane Array Detectors

In FT-IR spectroscopy, the amount of light that is transmitted from the sample to the detector is conventionally measured to generate the spectrum. Focal plane arrays (FPAs) can be employed as detectors in FT-IR spectroscopy. As suggested by its name, FPAs are aligned at the focal plane of the spectrograph (Salzer and Siesler, 2009). FPAs are capable of detecting photons of the incoming beam and thus are generally made of a semi-conductive material. Common materials used for the construction of FPAs include

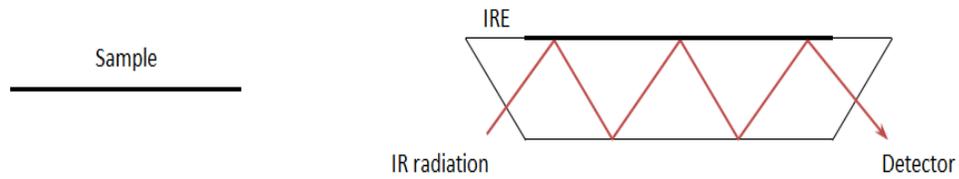
mercury-cadmium-telluride (MCT), indium-antimonide (InSb), silicon (Si), and germanium (Ge). The FPA material, regardless of type, is divided into an array of detectors; the number, size, and spacing of these detectors determine the resolution of the FPA (Global Security, 2010).

Utilizing FPA in conjunction with FT-IR spectroscopy is commonly referred to as a “multiplex approach” since FPA enables spectra of all pixels to be collected, and the interferometer of FT-IR spectroscopy enables spectra of all frequencies to be collected (Lewis et al., 1995). However, the size of the FPA can be a limiting factor in two ways. The first is known as “detector size-limited” in which the detected wavelength has a larger range than the FPA itself. The second is known as “diffraction-limited” and describes a situation in which the detected wavelength has a smaller length than the FPA detector (Salzer and Siesler, 2009).

2.3.2 Attenuated Total Reflectance

There are several accessories that can be coupled with FT-IR spectroscopy which alter the IR beam prior to its interaction with the sample. A commonly used FT-IR accessory is attenuated total reflectance (ATR). ATR employs a crystal which internally reflects the beam a set number of times, prior to the beam striking a detector (Fig. 2.4).

Figure 2.4 Attenuated total reflectance. *Source:* Martin-Gil et al., 2007

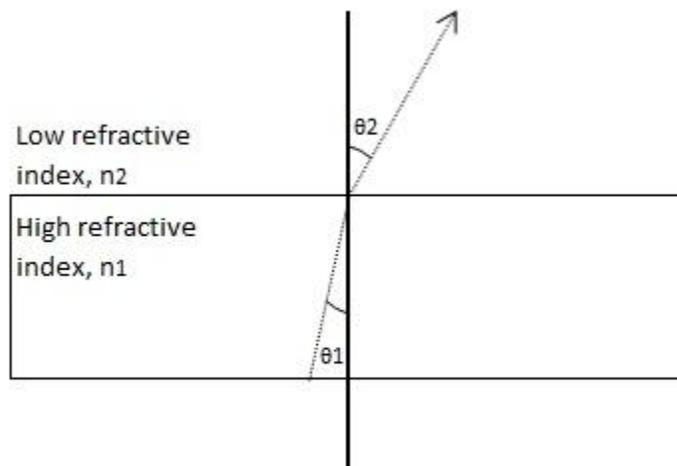


The concept of ATR is based on Snell's law (Fig. 2.5):

$$n_1 \sin \theta_1 = n_2 \sin \theta_2 \quad (2.22)$$

where n_1 and n_2 are the refractive indices of the two media in question (no unit), and θ_1 and θ_2 are the angles that the incident and refracted beam make with the normal (degrees).

Figure 2.5 Snell's law. *Source:* Griffiths and de Haseth, 2007



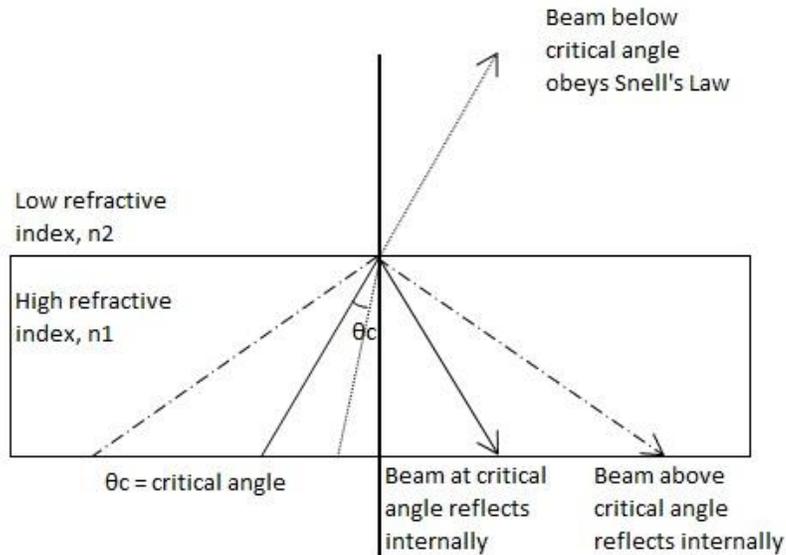
If medium 1 is significantly denser than medium 2, the angle of incidence in medium one (θ_1) will always be smaller than that of medium two (θ_2). If θ_1 is increased to a critical value then a condition known as total internal reflection will occur. This is when the angle of refraction of the incident beam in the dense medium is equal to 90° ($\theta_2 = 90^\circ$). The angle of incidence at which this occurs is known as the critical angle, θ_C , and is given by the equation:

$$\theta_C = \sin^{-1} \frac{n_2}{n_1} \quad (2.23)$$

The critical angle can continue to be increased and total internal reflection will occur within the dense medium; the reflected angle being equal to the angle of incidence (Fig. 2.6). Any internal reflection element (more commonly known as a crystal) employed in ATR must be extremely dense, possessing a high refractive index and critical angle in order to allow the crystal to completely reflect the incoming IR light through its medium (Griffiths and de Haseth, 2007). Common materials used to make ATR crystals include zinc selenide, germanium, and diamond (Perkin Elmer, 2005).

Figure 2.6 Total internal reflection and Snell's law. *Source:* Griffiths and de Haseth,

2007



ATR was developed in order to overcome several difficulties involved with sample preparation. The state of the sample can remain in its “as is” condition, as the IR beam is not required to completely penetrate the sample. This eliminates any potential changes in spectral characteristics during sample preparation such as grinding or dilution (Griffiths and de Haseth, 2007). Additionally, reproducibility of results, expense, and time consumption are improved when compared to conventional techniques (Perkin Elmer, 2005).

When utilizing ATR technology, the electric portion of the incoming electromagnetic IR beam is able to penetrate a few microns (0.5-5 μm) into the sample (Fig. 2.7). This is achieved at the point of contact between the crystal and the sample. The portion of the electric field that extends into the sample is known as the evanescent wave.

The strength of this evanescent wave decays as it reaches further into the sample, and consequently is a function of distance from the surface of the crystal (z):

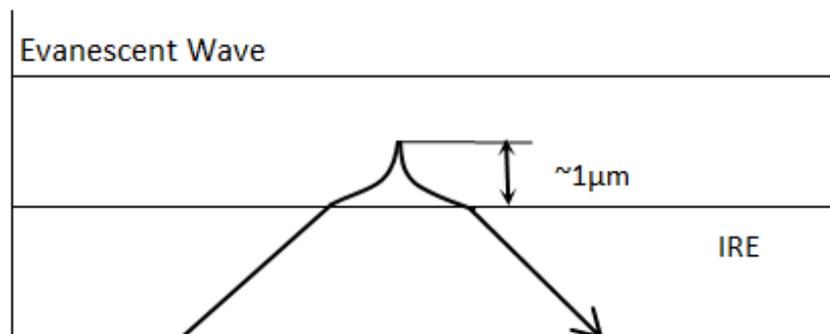
$$E = E_0 e^{-\gamma z} \quad (2.24)$$

where E is the strength of the electric field at distance z , E_0 is the initial strength of the electric field, and γ is a constant (Griffiths and de Haseth, 2007). The distance that the electric field penetrates into the sample is directly related to the wavelength of the IR beam. This depth of penetration can be calculated by:

$$D_p = \frac{\lambda}{2\pi n_1 \left(\sin^2 \theta_c - \left(\frac{n_2}{n_1} \right)^2 \right)^{1/2}} \quad (2.25)$$

where λ is the wavelength of the incoming light (cm) (Pike Technologies, 2000). The sample itself will absorb certain portions of the incoming IR spectrum, depending on its chemical properties. This absorption of IR energy by the sample will alter the IR signal as it is passed back through the crystal and onto the detector (Perkin Elmer, 2005). The sampling process occurs in the same way as without the ATR accessory (Pike Technologies, 2000), meaning additional user training is not required.

Figure 2.7 Evanescent wave properties. *Source: Martin-Gil et al., 2007*



Despite the immense benefits of ATR, several limitations do exist. The primary constraint with this method is that there must be sufficient contact with the ATR crystal and the sample in order to obtain quality spectra (Griffiths and de Haseth, 2007). This can become difficult or impossible, depending on the nature of the sample. The ATR crystal itself is fragile and subject to scratching, and caution must be exercised when using the accessory. Due to the fact that the sample is making contact with the crystal, the crystal must also be thoroughly cleansed with a solvent after every scan to ensure that data from one sample does not affect the data obtained from a subsequent scan. The type of sample that can be scanned by ATR is also limited. Abrasive and corrosive samples can scratch and degrade the crystal, respectively, and should be avoided (Pike Technologies, 2000). The index of refraction of the sample must also be known and, in some situations, this can be difficult or impossible to obtain. Additionally, some solids exhibit weak spectra when they are confined to a very small contact area of the crystal. Finally, ATR spectra require a correction factor and subsequent calculation in order to relate them to spectra collected by the conventional technique; however, this is normally computed via computer software and thus is not tedious to the user (Perkin Elmer, 2005).

2.4 Instrumentation:

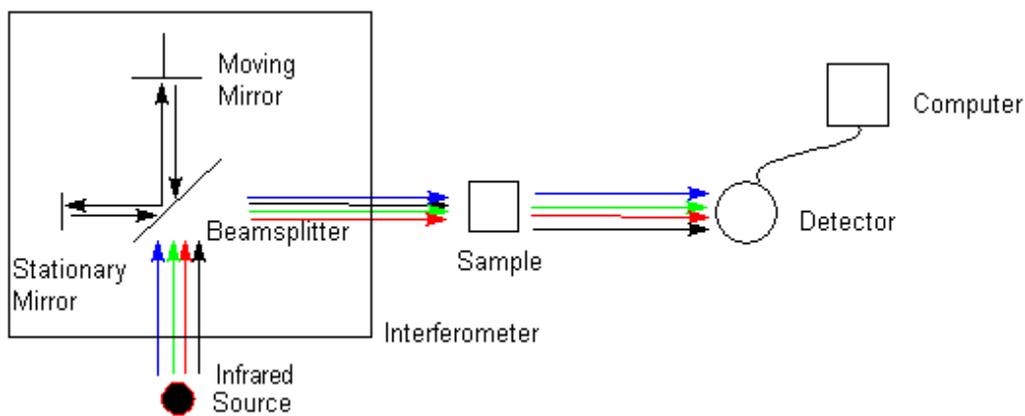
Infrared spectrometers have been in existence for practical applications for almost a century. These instruments were originally of a dispersive type, in which frequencies of the infrared light source are separated via a grating apparatus. There are several complications that exist with the dispersive spectrometers, namely a slow scanning speed and mechanical fatigue of the equipment due to moving components in the spectrometer.

FT-IR spectrometers are capable of overcoming the two obstacles faced by dispersive infrared spectrometers. Scanning speed is accelerated via an apparatus known as an interferometer which is capable of measuring all infrared frequencies simultaneously. Mechanical fatigue of the equipment is reduced by using a beamsplitter, which only incorporates one moving component in the spectrometer. In addition to overcoming these two obstacles, the instrumentation employed by FT-IR spectrometers allow spectral readings that have a high SNR and accurate wavelength precision (Griffiths and de Haseth, 2007).

2.4.1 Overview of FT-IR Spectrometers

FT-IR spectrometers consist of five separate components: an infrared (IR) light source, an interferometer, a sample, a detector, and a computer (Fig. 2.8).

Figure 2.8 Schematic of the components of a FT-IR spectrometer. *Source:* Gable, K (2000)



IR light is emitted from a blackbody source and focused through a slit which regulates the amount of energy intensity passed onto the interferometer. Light passes through the interferometer, where beam splitting occurs, and then shines onto the sample.

The beam will either be absorbed or transmitted by the sample. This absorbance and transmittance of energy is what essentially creates the unique molecular fingerprint of the sample in question. The part of the beam that is transmitted by the sample passes on to the detector. The detector measures the resulting interferogram signal that has passed through the sample and relays this information to a computer where the Fast Fourier Transform (FFT) is applied and the interferogram is transformed into a spectrum (Thermo Nicolet Corporation, 2010).

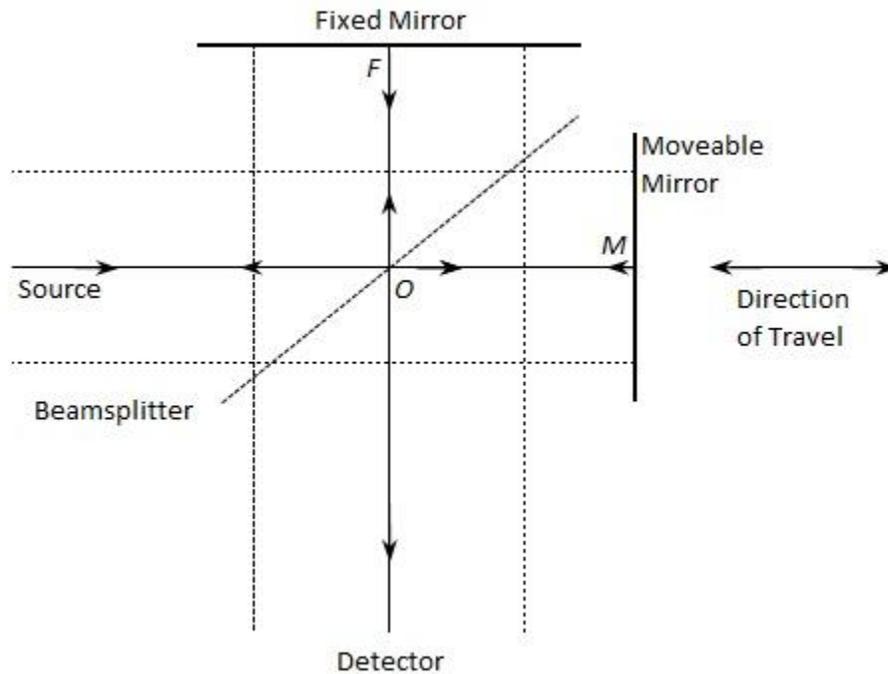
A background spectrum must be obtained prior to placing any substance in the sample compartment. This is in order to gather spectral information of the instrumentation itself, and the environmental conditions of where the instrument is being housed. Once a sample is scanned, the background spectrum enables a percent transmittance to be calculated and the resulting spectrogram is free of any instrumental and environmental characteristics (Thermo Nicolet Corporation, 2010).

2.4.2 Interferometer:

The interferometer is a unique optical device that is capable of measuring all infrared frequencies simultaneously. Due to this quality, it significantly expedites the scanning process involved with FT-IR spectroscopy (Thermo Nicolet Corporation, 2010). The design of all interferometers came from the foundation provided by the Michelson interferometer which was created in 1891. The basic theory behind all types of interferometers can be interpreted by gaining an understanding of the operating principle of the Michelson interferometer (Griffiths and de Haseth, 2007).

The Michelson interferometer operates based on the principle of interference. It employs an optical device, known as a beamsplitter, which is oriented at a 45° angle to the incoming IR light source. This orientation allows the incoming light to be divided into two beams at right angles to one another. The beamsplitter is generally constructed of a non-absorbent material and is capable of reflecting and transmitting 50% of light to a fixed mirror and 50% of light to a mirror which is able to move (Fig 2.9) (Åmand and Tullin, 1999). The moving mirror can be moving at a constant speed (continuous interferometer) or at a rapid velocity between equally spaced points of zero velocity (step-scan interferometer). After the beams are reflected from their respective mirrors, they are recombined at the beamsplitter and are sent out of the interferometer, perpendicular to the direction of the incoming IR light source. The intensity of the beam emerging from the interferometer varies, dependent upon the path difference that the beams must travel between the two mirrors (Griffiths and de Haset, 2007).

Figure 2.9 Michelson interferometer. *Source:* Griffiths and de Haseth, 2007



The path difference between these two beams is found by the position of the two mirrors and is known as optical retardation:

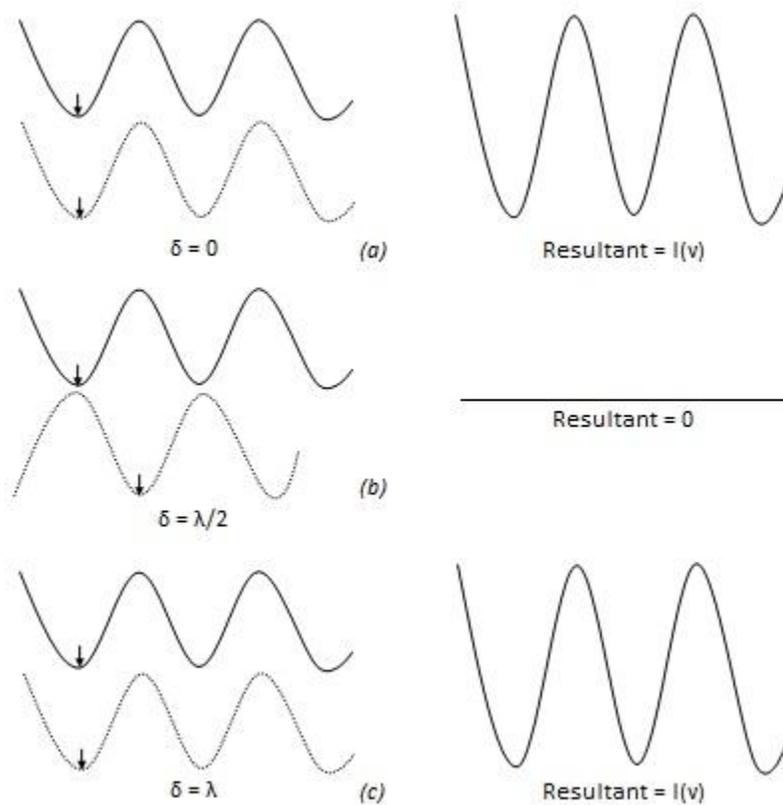
$$\delta = 2(OM - OF) \tag{2.26}$$

where OM is the distance between the beamsplitter and the moving mirror (cm) and OF is the distance between the beamsplitter and the fixed mirror (cm). From this equation, it can be seen that if the distance between the beamsplitter and the fixed and moveable mirrors are equal, there will be no retardation between the beams. The two beams are perfectly in phase and constructive interference will occur, causing all source energy to be emitted from the interferometer. Alternatively, if a retardation of 0.5 exists between the two mirrors, the beams will be exactly out of phase and exhibit destructive interference; therefore, none of the source energy will exit the interferometer, all light

returning to the source. The condition of constructive interference is restored again when an optical retardation of $\delta = \lambda_0$, where λ_0 is the wavelength of the source beam, exists.

Fig. 2.10 displays how the phase difference between the fixed (solid line) and moving (dashed line) mirrors affects the resultant beam exiting the interferometer (Griffiths and de Haset, 2007). Hence, as the moveable mirror moves back and forth in position, light and dark bands, corresponding to constructive and destructive interference, are given in the output (Åmand and Tullin, 1999).

Figure 2.10 Phase difference between the fixed (solid line) and moveable (dashed line) mirrors and resultant output from the interferogram. *Source:* Griffiths and de Haset, 2007



2.4.3 Generation of an Interferogram

The intensity of the beam that is picked up by the detector is a function of optical retardation and is denoted by the symbol, $I(\delta)$. The beam is composed of a direct current (DC) and an alternating current (AC) component, but only the AC component of the signal is used to generate the interferogram. The interferogram is given by the equation:

$$I(\delta) = 0.5I(\tilde{\nu}_0)\cos 2\pi\tilde{\nu}_0\delta \quad (2.27)$$

where $\tilde{\nu}_0$ is the wavenumber (cm^{-1}) of the source beam and $I(\tilde{\nu}_0)$ is the intensity of the source beam (Griffiths and de Haseth, 2007).

There are several factors that affect the quality and amplitude of the signal measured by the detector. In practice, it is unlikely that any beamsplitter will perfectly reflect and transmit 50% of the incoming light to each mirror. The offset from the ideal 50% target can be compensated for by a factor known as the beamsplitter efficiency. The beamsplitter efficiency is multiplied to Eq. 2.27 in order to generate an accurate interferogram that is representative of both mirrors receiving equal portions of the beam. In addition to disproportionate beam splitting, not all detectors will have the same response to the different infrared frequencies found in the striking beam. In order to correct this, there are a number of filters that can be used to eliminate signals that are outside of the spectral range of interest (Griffiths and de Haseth, 2007), as well as a number of correction factors that can help with the responsivity of the detector. Since the quality and amplitude of the interferogram is dependent upon the beamsplitter efficiency and the detector responsivity, Eq. 2.27 can be modified to:

$$S(\delta) = B(\tilde{\nu}_0) \cos 2\pi \tilde{\nu}_0 \delta \quad (2.28)$$

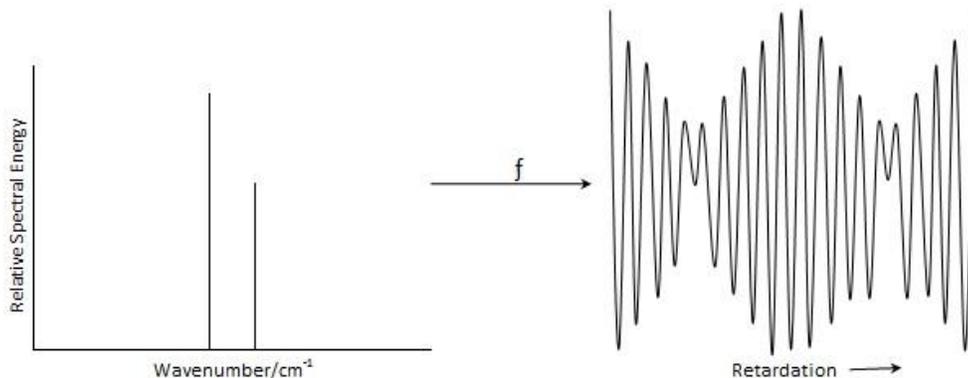
where $S(\delta)$ is the AC component of the signal (volts) and $B(\tilde{\nu}_0)$ is the intensity of the source beam, as modified by the aforementioned instrumental characteristics (Griffiths and de Haseth, 2007).

Once the interferogram signal is captured by the detector, it is passed onto a computer that applies the FFT algorithm. Due to the fact that IR light is composed of radiation spanning numerous wavelength values, the final interferogram of the FT-IR technique incorporates a multitude of interferograms, each corresponding to one wavelength value. Since the IR source is a continuum of wavelengths, the resulting interferogram must be represented by the integral equation:

$$B(\tilde{\nu}) = 2 \int_0^{+\infty} S(\delta) \cos 2\pi \tilde{\nu} \delta d\delta \quad (2.29)$$

The above equation accounts for the spectral range over which the scan was taken. In practice, this spectral range is of a finite resolution; this arises from the fact that the optical retardation is limited by the size of the spectrometer, and cannot possibly vary between 0 and $+\infty$ cm (Griffiths and de Haseth, 2007). Fig. 2.11 shows an interferogram and its corresponding spectrum. The interferograms are measured in terms of optical retardation, whereas the spectra are measured in terms of wavenumber.

Figure 2.11 Interferogram and corresponding spectrum. *Source:* Griffiths and de Haseth, 2007



2.5 Chemometrics

Chemometrics is a discipline of chemistry that applies mathematical and statistical approaches to data in order to improve chemical measurements and extract useful information (Liu, et al., 2010). Chemometrics involves a variety of techniques; the method of choice depends on the quality of spectra obtained and the desired outcome of the project. There are some instances where the data collected from FT-IR spectroscopy is impossible to analyze without chemometric methods. Therefore, tools for extracting important, interpretable information from the measurements are required (Salzer and Siesler, 2009).

Chemometric techniques that are used in conjunction with the interpretation of FT-IR spectra include both pre-processing and quantitative analysis approaches. Pre-processing techniques encompass practices such as de-noising, baseline corrections, data compression, and light scattering correction. Common quantitative analysis approaches

include multiple linear regression, principal component analysis, principal component regression, and partial least squares regression.

2.5.1 Pre-processing Methods:

The spectra obtained from FT-IR analyses can be subject to a variety of pre-processing methods prior to analyzing the results of the scan. Although the spectrum contains significant useful information about the sample in question, there are several common issues that can affect the quality of the spectra that may need to be removed. Spectral images may contain a high noise level, irregularly shaped baseline contributions, outlier data, and light scattering effects that can affect the quality of the spectrum and may need to be eliminated prior to using the spectrum for data analysis (Salzer and Siesler, 2009).

De-noising: There are several methods of de-noising that currently exist in order to improve the SNR of a spectrum. The primary approaches include smoothing procedures, signal filtering, and decomposition of the original data set.

Smoothing procedures remove all high frequency components, regardless of their amplitude, from the spectrum (Leung et al., 1998). Once an interferogram is collected it must be truncated due to the finite optical path of the moving mirror. There are several apodization functions, such as the Happ-Genzel, Norton-Bier, or Gaussian, that smooth the spectrum produced from the truncated interferogram by reducing the effect of the sidelobes. This ultimately improves the SNR by convoluting the acquired spectrum (Griffiths and de Haseth, 2007). However, the most commonly used function for smoothing procedures is the Savitzky-Golay method. This approach fits the spectrum to a least squares polynomial function to effectively smooth the signal without distorting its

relative distribution (Savitsky and Golay, 1964). The downside to implementing a smoothing technique with FT-IR spectroscopy is that the spectral bands generally become larger and spectral resolution is lost. It is recommended that sampling initially be done at a lower noise level rather than performing a smoothing function on the obtained spectrum (Griffiths and de Haseth, 2007).

Signal filtering eliminates noise from the spectrum by removing all high frequency signals associated with it (Yeniay and Götaş, 2002). The Kalman filter is effective at reducing the Gaussian noise of a signal. It operates based on the principle of recursion, in that it generates least squares estimates of observations in an attempt to produce data values that are closer to true measured values, thus eliminating noise from the spectrum (Kleeman, 1996).

Decomposing the original data set into variables that estimate the relevant information contained in a spectrum is another common method of de-noising. Mathematical techniques, such as principal component analysis and partial least squares regression, can be employed to create a new data set that is free of redundant information or noise by decomposing the old one. These methods are discussed in further detail in Section 2.5.2 of this report.

Baseline Correction: Ideally, the baseline of a spectrum will be a flat line at 100% transmittance (0% absorbance). However, once a sample is scanned, there are many factors such as scattering, reflection, temperature, or instrumental characteristics, which can alter the spectral baseline and detrimentally affect the obtained results. These must be corrected in order to achieve optimal spectral data (Griffiths and de Haseth,

2007). Corrections made to the baseline are best performed after a de-noising technique has been applied to the image. Baseline corrections are based upon mathematical functions and work with derivative spectra as opposed to raw spectra (Salzer and Siesler, 2009).

A common form of baseline error is an offset. This type of error affects all spectral points, among all wavelengths. Generally, the absorbance of the spectrum is raised in comparison to the background spectrum. This may occur as a result of light attenuation by the sample. An adjustment technique known as a one-point baseline correction fixes baseline offset. This method uses a constant, which is typically taken as the minimum absorbance value, and subtracts that constant from all spectral points (Griffiths and de Haseth, 2007).

Another correction technique, known as two-point baseline correction, is employed when the baseline is sloping in a linear fashion, from one end of the spectrum to the other. This correction method uses interpolation through two selected points and subtracts the resulting line from the spectrum. Two-point baseline corrections can be extended to correct nonlinear baseline slopes by selecting several points and employing a series of linear interpolations. The resulting lines from each section of the linear interpolations are subtracted from the entire spectrum. Caution must be exercised when using this correction technique, as discontinuities in the baseline slope may result. Baseline slopes that are nonlinear are better corrected by fitting a nonlinear correction function, such as a high-order polynomial, to the data and subtracting that result from the spectrum (Griffiths and de Haseth, 2007).

Baseline distortions may also arise from resonant Mie scattering (Life Science Weekly, 2012). Resonant Mie scattering (RMieS) correction has recently been found to be effective at removing baseline irregularities. Bassan et al. (2009) developed a modification to the traditional extended multiplicative signal correction which has been successful in removing baseline bulges (Bassan et al., 2010a; Bassan et al., 2010b; Bambery et al., 2011). This technique models the spectrum as “a sum of scattering components, a sloped line and the desired absorption components”. The third term is achieved by estimating and subtracting the first and second terms from the spectrum (Martin et al., 2010).

The majority of spectral software packages come with functions that can automatically correct the baselines of the spectra obtained by the user. These automatic techniques are applied to all regions of the spectrum. In areas of the spectrum where there are no spectral peaks present, a linear or polynomial correction method is usually employed to obtain the desired baseline. If spectral peaks exist, corrected data from the adjacent regions are made to amend the baseline. Several other automatic baseline correction techniques are available. These include, but are not limited to, iterative stripping, baseline standard deviation, binning, and ultralow-frequency filtering (Griffiths and de Haseth, 2007). The interested reader can find detailed information about these techniques in other spectroscopic texts.

Data Compression: Data compression is often required when the size of spectral images is extremely large. This is primarily because computation times are inversely proportional to spectral size and a large data set results in a slow computation time. Spectral dimension compression, spatial dimension compression, or a combination of the

two can be used to achieve successful results. Reducing the number of pixels or spectral channels is another data compression technique that is regularly used. However, the most common technique employed to compress a spectral image is spectral binning (Salzer and Sielser, 2009)

Spectral binning minimizes spectral variations of peak positions by dividing the spectrum into multiple regions of equal size (Anderson et al., 2011) and adding data in intervals of neighbouring spectral channels or pixels to form a single measurement (Salzer and Siesler, 2009). Additionally, this method reduces the dimensionality of the spectrum, which leads to a data set that highlights only key spectral features, making the spectrum more suitable for multivariate statistical analyses such as principal component analysis. Despite these immense benefits, resolution is lowered. Additionally, binning may result in spectral shifts near the boundaries of the spectrum which can lead to errors in quantitative analysis. These shifts can be corrected for by weighting the peaks based on their distance from the center of the bin or by choosing the size and location of each bin in a dynamic, iterative manner (Anderson et al., 2011).

Light Scattering Correction: Light scattering effects are common with samples that are solid and opaque. Resonant mie scattering is usually the most pronounced light dispersion effect that occurs with IR spectra (Martin et al., 2010). Several light scattering correction techniques exist, the most common being resonant Mie scattering correction, standard normal variate (SNV) with detrending, multiplicative scatter correction (MSC), and wavelet transform (WT) (Wang and Paliwal, 2007).

Resonant Mie scattering doesn't predominately causes baseline errors. However, this light dispersion effect can also responsible for peak alterations such as distortion of peak shape and shifting of peak position (Life Science Weekly, 2012). Resonant Mie scattering correction techniques were previously discussed in the baseline correction section.

Barnes et al. (1993) describe the SNV transformation, as it applies to spectroscopy. SNV is a technique that is used on individual spectra, which centers the spectrum by scaling it with its own standard deviation. The transformed SNV spectral values can be obtained by the equation:

$$Y_{SNV,i} = \frac{[Y_{LOG,i} - mean(Y_{LOG,i})]}{SD_{LOG}} \quad i = 1, 2, \dots, k \quad (2.30)$$

where $Y_{LOG,i}$ is the i^{th} value of the raw data set (no unit) at k wavelength (no unit), $mean(Y_{LOG,i})$ is the mean value of the raw data set (no unit), and SD_{LOG} is the standard deviation of the raw data set (no unit).

Solely performing the SNV transformation is generally not sufficient enough to remove all errors associated with the spectrum, as irregularly shaped baselines are often still persistent. Detrending is performed to remove this effect. Detrending involves the use of a quadratic equation in which the spectral variables are computed as a function of wavelengths:

$$\hat{Y}_{SNV,i} = A_{SNV} + B_{SNV}W + C_{SNV}W^2 \quad (2.31)$$

where A , B , and C are regression coefficients, and W is the wavelength range (nm) from 1100-2498nm at 2nm intervals (Barnes et al., 1993). Since $\hat{Y}_{SNV,i}$ is an estimate of the

transformed spectral data, based on a regression equation, another calculation which compensates for the estimation must be performed:

$$Y = Y_{SNV,i} - \hat{Y}_{SNV,i} \quad (2.32)$$

The above equation results in data that has been subjected to a SNV transformation as well as detrending to correct any baseline irregularities (Barnes et al., 1993).

MSC is a transformation that compensates for multiplicative and additive effects found in spectral data. MSC is unique in that it deals with optical interference, rather than chemical interference that is often dealt with when using other transformations. Values describing the population of a particular data set are incorporated into the calculation and thus MSC models need to be established for every data set. The MSC transformation is given by the equation:

$$x_{ik}(new) = \frac{[x_{ik}(old) - a_i]}{b_i} \quad (2.33)$$

where $x_{ik}(old)$ is the optical value before the MSC transformation (no unit) of the sample i in k wavelengths, $x_{ik}(new)$ is the optical value after the transformation (no unit), a_i is an estimation of the specular reflection in the data (no unit), and $(\frac{1}{b_i})$ is an estimation of the scatter interferences in the sample (no unit). The constants a and b are estimated using the following least squares regression model for an individual spectrum:

$$x_{ik} = a_i + b_i X_{..} + e_{ik} \quad (2.34)$$

where X_i is the average optical sample of the calibration set (no unit), and e_{ik} is the error of the model (no unit). Estimations of a_i and b_i need to be done for each sample, using k spectral measurements (Maleki et al., 2006).

WT is an efficient, high speed function that is able to preserve both time and frequency information of the signal. WT has been termed a “mathematical microscope” as it is capable of focusing on a small portion of a signal. Leung et al. (1998) detail several wavelength transform functions that can be employed under various data conditions; however, all basis functions for WT can be derived from:

$$\Psi_{a,b}(\lambda) = a^{-1/2} \Psi\left(\frac{\lambda-b}{a}\right) \quad a, b \in R \text{ and } a \neq 0 \quad (2.35)$$

where $\Psi(\lambda)$ is the mother wavelet (no unit), and a and b are the scale and position parameters, respectively, in real number, R (no unit).

Multiple reviews (Walczak et al., 1983; Leung et al., 1998; Azzouz et al., 2003; Wang and Paliwal, 2006) have been done to compare and contrast these different light scattering correction techniques in an attempt to determine the most advantageous approach. The basic outcome of these reviews is that the performance of each method is dependent upon the data set, the applications, and the desired results of the experiment.

2.5.2 Quantitative Analysis

There are multiple quantitative analyses that can be performed in order to extract useful information from a given spectrum. The most common approaches for FT-IR spectroscopy are Multiple Linear Regression (MLR), Principal Component Analysis

(PCA), Principal Component Regression (PCR), and Partial Least Squares Regression (PLSR) and are highlighted in this review.

Multiple Linear Regression: MLR is an effective technique for data sets that are not large or redundant, and for data in which the relationship between the data points can be readily extracted. MLR seeks to provide the effect of a number of predictor variables (dependent variables) on a single outcome (independent variable) (Marill, 2003; Benoudjit et al., 2009).

MLR uses least squares equations to predict the regression coefficients of a model in order to provide a linear relationship between the dependent and independent variables in the data set. Once predicted, this model takes the form:

$$Y = \beta_0 + \beta_1x_1 + \beta_2x_2 \dots \beta_kx_k + e \quad (2.36)$$

where Y is the independent variable (no unit), β_0 to β_j are the regression coefficients (no unit), x_1 to x_k are the predictor variables (no unit), and e is the error term associated with the model (no unit). With a sample size of n using k predictor variables and r response variables, two matrices are employed to predict the regression coefficients: a predictor or explanatory matrix (X) of size k by n and a response matrix (Y) of size r by n . The following equation is used:

$$Y = X\beta + E \quad (\beta = b_0, b_1, \dots b_k) \quad (2.37)$$

where β is the regression coefficient matrix of size 1 by $k+1$, and E is the error residual matrix of size 1 by n . All matrices have no units associated with them (Mahmoud, 2008).

Caution must be exercised when using MLR if the data set is large, redundant, or has a high correlation between two or more independent variables within the model. In these situations, MLR will prove to be ineffective at describing the data and other methods of quantitative analysis should be sought (Fletcher, 2008).

Principal Component Analysis: PCA is an iterative technique that can be used when there is a degree of redundancy in the data set (Guo et al., 2009). PCA highlights similarities and differences between patterns within a given data set. This technique is accomplished by transforming the data, and essentially reducing the number of dimensions associated with it, in order to give a data set that readily explains trends within the measured data (Griffiths and de Haseth, 2007).

The original data matrix (X) is essentially decomposed into two matrices. The score matrix for X , T , of size n by e , and the loading matrix for X , P , of size k by e , which can be described by the equation:

$$X = TP^T + E \quad (2.38)$$

where E is the residual matrix for X of size n by k . All matrices have no units associated with them, n and k are the same variables as previously defined with MLR, and e is the eigenvalue for the matrix. The score matrix is a projection of the original data onto the new coordinate system and gives information about which image coordinates are most relevant to the overall data analysis. The loading matrix gives the values of the weights of each variable used with computing the principle component and indicates relevant spectral features within the image (Salzer and Siesler, 2009).

The transformation occurs by determining which principal component (PC) will be removed first from the calibration spectra (multiple algorithms will calculate this). Once the first PC is removed, only the residuals are left in the calibration set. The largest PC of the residuals (the second principal component) is then calculated and removed from the data set, leaving a set of second residuals. The process is repeated until the number of PCs is equal to the number of components in the calibration set (Griffiths and de Haseth, 2007).

Not all of the calculated PCs in a data set will be significant. As the PCs are determined and removed from the data set, the contribution of each PC to the calibration data set is recorded. These contributions are known as the eigenvalues and are used as a measure of significance of the PCs. The PCs that are considered insignificant for the data set can be regarded as noise (Griffiths and de Haseth, 2007).

The final data set gives rise to a transformed spectrum, with minimal redundant information, that can be interpreted more easily than the initial data set (Salzer and Siesler, 2009). Two quantitative methods incorporate the PCA approach into their techniques. These are: principal component regression and partial least squares regression.

Principal Component Regression: PCR utilizes the basics of PCA but attempts to construct factors that are regressed with the *Y* variable by utilizing a two-step process (Griffiths and de Haseth, 2007). The first portion of the process is using PCA to determine the principal components, and the second is to regress the components against the scores. Essentially, PCR uses the score matrix in place of the original data matrix:

$$T = XP \tag{2.40}$$

$$Y = TB + E \tag{2.41}$$

This relationship allows the principal components of X to be used as regressors on Y (Abdi, 2003). PCR analysis eliminates problems associated with multicollinearity (Abdi, 2003) because the score vectors within the matrix are orthogonal to one another (Griffiths and de Haseth, 2007). However, the final result of PCR may incorporate more principal components and scores than are needed, and these principal components may not be correlated to the concentrations of the constituents of the sample. Partial least squares regression is a technique that compensates for these shortcomings of PCR by decomposing both the absorbance and concentration matrices (Griffiths and de Haseth, 2007).

Partial Least Squares Regression: PLSR is an appropriate technique to use when the independent variables of a data set are highly correlated and the data set consists of a large number of independent variables in comparison to the amount of dependent variables present (Abdi, 2003). PLSR is used to construct predictive models and thus should not be used to decipher a relationship between two variables. PLSR combines features from both PCR and MLR.

PLSR is similar to PCR in that it attempts to construct factors regressed with Y (Tobias, 1999). However, instead of solely using the decomposition of X to predict new factors, PLSR decomposes X and Y together and then uses a regression step to predict Y (Abdi, 2003). The simultaneous decomposition is the primary feature that makes this technique powerful, robust, and more complex than the aforementioned quantitative

methods (Griffiths and de Haseth, 2007). The independent variables (X) are decomposed using Eq. 2.38 of PCA. The dependent variables (Y) are decomposed by:

$$Y = Uq^T + F \quad (2.42)$$

where U is the score matrix for Y of size n by e , q is the loading vector of 1's for Y of size 1 by e , and F is the error matrix for Y of size n by 1. All matrices have no units associated with them, and n , k and e are the same variables as previously defined with PCA (Doymaz et al., 2003).

There are two different kinds of PLSR algorithms, denoted as PLSR-1 and PLSR-2. PLSR-1 calculates the scores and loadings independently, whereas PLSR-2 iteratively calculates the loads and scorings at the same time, using only one set of loadings for all calculations. PLSR-2 is a much faster method than PLSR-1; however, PLSR-1 is generally more accurate due to separate calculations (Griffiths and de Haseth, 2007).

PLSR has many advantages over the other quantitative techniques. This method is capable of covering the full spectral range, improving the precision of the data, modelling specific sample components, and analyzing complex samples. PLSR is a slower method compared to the aforementioned ones, and requires more mixture spectra than constituent spectra in order to give accurate results (Griffiths and de Haseth, 2007).

Yeniay and Göktaş (2002) compared several regression methods, including PCR and PLSR in order to determine which analysis yields the best results when applied to a data set with high collinearity. It was found that among all of the methods tested, PLSR had the “highest predictive ability with the smallest number of factors” available within the model.

Another comparison of regression techniques was done by Geladi and Kowalski (1986). They compared MLR, PCR, and PLSR regression techniques and developed an algorithm for PLSR. The comparison highlights the problems MLR has when collinearity exists between the variables in the model. MLR provides a model that fits the data, but fails to predict future observations (Tobias, 1999). PCR addresses this shortcoming of MLR by decomposing data via the PCA method and then constructing factors via regression. However, PCR doesn't consider all relevant data found in the original set; only the most dominant dimensions are included in the final analysis. PLSR builds upon PCR by decomposing variables in both the X and the Y dimensions and emphasizing the correlation between these two variables (Qin, 1993).

2.6 FT-IR for Fusarium Detection

There are many challenges in the agricultural industry associated with the detection of Fusarium in wheat, the primary difficulty being a lack of an objective detection method. FT-IR spectroscopy has only recently been explored as a detection technique for identifying Fusarium in grain commodities. A brief review of the implications of Fusarium infection in wheat, as well as the recent applications of FT-IR spectroscopy as it relates to Fusarium detection, is given. Recommendations for future work based on previous research are also provided.

2.6.1 Fusarium Head Blight

Fusarium head blight (FHB), caused by the filamentous fungus *Fusarium* (several species exist), is a fungal disease that infects grain commodities around the world. The main commodities affected are: wheat, barley, and corn. *Fusarium* spp. was first discovered in

1884, with the first recorded appearance of the fungus in Canada in 1923. However, it was not until the 1980s that the disease became prevalent in the Canadian Prairies. Since then, *Fusarium* spp. has been spreading throughout the country, causing increasing proportions of Fusarium damaged kernels (FDK). FDKs are harmful to the Canadian economy as they cause losses in grain quality and yield (Clear and Patrick, 2010), as well as lead to rejection and downgrading of the grain commodity (Osborne and Stein, 2007). As a comparison, losses resulting from FHB in three states just south of Canada (North Dakota, South Dakota, and Minnesota), exceeded \$1 billion USD in the 1993 fiscal year (De Wolf et al., 2003).

Fusarium is a pre-harvest fungus and is therefore only capable of infecting grain kernels prior to harvest. The life cycle of the FHB diseases is as follows: source, dispersal of inoculation, and infection (Xu and Berrie, 2005). The agents that cause FHB (source) survive and grow on and within plant tissue residues such as small grain stems and roots. The fungi survive as mycelium on the aforementioned residues during the fall, winter, and spring. In the spring, the fungi produce abundant mycelium, as well as spores, and inoculate the grain kernel under favourable fungal conditions. The presence of *Fusarium* spp. is often associated with high moisture conditions (relative humidities greater than 90%) during flowering, for extended time periods of two to three days. Despite this generality, infection can still occur if high moisture conditions don't exist (De Wolf et al., 2003) and beyond the flowering stage, up until the grain is harvested (Shahin and Symons, 2011). The pathogen is dispersed by both short- and long-distance dispersal modes. The amount of crop residue left behind on the soil from the previous year's crop is the most significant form of short-distance dispersion of FHB. Long-distance dispersal

modes include movement by wind, rain, and insects (Osborne and Stein, 2007). The fungus typically infects the wheat plant's cells in a subcuticular fashion and the interior and exterior of the leaf stomata (Xu and Berrie, 2005).

FHB causes visibly damaged seeds (FDKs), as well as seeds that are visually asymptomatic. FDKs generally have a shrivelled and white appearance to them and may be accompanied by a pink or purple-black fungus (Osborne and Stein, 2007). However, there are some cases in which FDKs are difficult to decipher from regular grain kernels (visually asymptomatic). For example, Fig 2.12 depicts FDKs which were artificially infected at different life stages, the earliest lifestage infection on the far left hand side, and the latest lifestage infection on the far right hand side. The four kernels on the left hand side are visibly infected, whereas the two kernels on the right hand side appear healthy, even though they contain the fungus (Clear and Patrick, 2010). Fig 2.12 indicates that infection at early lifestages causes the most visible damage, as well as the most mycotoxin production, to the seed (Shahin and Symons, 2011), but that differentiation between uninfected and FDKs can be difficult with the naked eye.

Figure 2.12 Fusarium Damaged Kernels (FDK) artificially infected at different life stages. *Source:* Clear and Patrick, 2010



There are six well documented species of *Fusarium* (Desjardins et al., 1993). Four of these are found in North America and are capable of causing FHB: *F. graminearum*, *F. culmorum*, *F. avenaceum*, and *F. crookwellense*. In Canada, and much of the world, *F. graminearum* is the most predominant species of *Fusarium* that causes FHB. This is due to its high genetic variability, tolerance to different temperature and moisture conditions, aggressive nature (Osborne and Stein, 2007), and its ability to produce several different mycotoxins, such as deoxynivalenol (Clear and Patrick, 2010).

Grain that is classified as FHB may contain mycotoxins. However, the occurrence of mycotoxins depends on several factors such as environment, the species of *Fusarium* infecting the grain, how long the grain has been infected for, and the crop variety (McMullen et al., 2008). An extensive biosynthetic pathway explanation for production of trichothecenes from *Fusarium* spp. has been documented by Desjardins et al. (1993). The interested reader is referred to their work.

The major mycotoxins produced by *Fusarium* spp. are: deoxynivalenol (DON), nivalenol (NIV), zearalenone (ZEA), and moniliformin (MON). The mycotoxin DON, which is sometimes referred to as vomitoxin, is the most commonly detected mycotoxin throughout the world. DON belongs to the family of trichothecenes, and is the least toxic member. Despite DON's low toxicity, being the most commonly detected vomitoxin throughout the world allows it to serve as an indicator for the presence of more harmful trichothecenes. *F. graminearum* and *F. culmorum* are the two species of *Fusarium* that are capable of producing DON. Consumption of DON by livestock causes symptoms of weight loss, diarrhoea, vomiting (Abramovic, et al., 2007), immunosuppression, and inhibition of protein biosynthesis (Biselli et al., 2005). If ingested by humans, DON can

cause inflammation of tissues and eyes, as well as symptoms similar to that of livestock (Nie, et al., 2007). Furthermore, flour colour, baking quality, and ash content are compromised by the presence of DON (Shahin and Symons, 2011). In Canada, the maximum permissible level of DON in food is 2.00 ppm of grain (Abramovic et al., 2007).

There is no portion of the wheat kernel that is resistant to infection by *Fusarium* spp. (Pinson-Gadais et al., 2007). A study by Betchel et al. (1985) on the effects of *F. graminearum* in wheat kernels found that the fungus destroyed the starch granules, storage proteins, and cell walls of the kernel. The fungus was found to initially invade the pericarp and aleurone of the kernel, and then migrate towards and enter the starchy endosperm by degrading the cell walls. The fungus was found to disperse itself throughout the endosperm where it digested the protein and starch of the kernel.

Boyacioglu and Hettiarachchy (1995) analyzed wheat kernels infected by *F. graminearum* for changes in their biological characteristics after infection by the fungus. They analyzed the kernels in terms of carbohydrates, proteins, and lipid content. The study found that there was an increase in the total protein and lipid content of the kernel with a decrease in the cellulose content. The authors also observed that this fungus attacks the pericarp and aleurone parts of the kernel first, and enters the endosperm soon afterward by penetrating the cell wall. These results confirm what was found by Betchel et al. (1985).

Kang and Buchenauer (2002) studied the effects of *F. culmorum* on wheat kernels and found that the fungi caused degradation to numerous areas of the wheat kernel cell,

such as the cytoplasm, cell organelles, and parenchyma cells. Additionally, disintegration of the cell wall and pronounced alteration of the middle lamella matrix was observed. The degree of degradation experienced by the cell wall is proportional to the proximity of hyphae, with areas in contact with the hyphae undergoing the greatest reduction. The authors hypothesized that the degradation of the cell wall and middle lamella matrix is the result of a synchronized release of different deteriorating enzymes from the fungal species.

When trichothecene toxins such as DON are produced, they are generally localized in the cell wall, epidermis, and cytoplasm, which are the outermost portions of the wheat kernel (Pinson-Gadais et al., 2007). However, DON is a water soluble toxin and is capable of migrating to all portions of the wheat kernel. Due to this property, if DON is discovered in any part of the kernel, it is likely the entire kernel will need to be discarded (Sinha and Savard, 1997). Trichothecene toxins cause adverse effects to cell functions by inhibiting protein, DNA, and RNA synthesis, convoluting the cytoplasm and diminishing cell organelles (Kang and Buchenaur, 2000).

Currently, in the Canadian grading system, a sample of the grain product is visually inspected for *Fusarium* spp. infection, and if present, the sample is sent to a laboratory where testing for the presence of DON occurs. This process is tedious, time-consuming, and subjective. When FDKs are only visible as slightly damaged, this process is even lengthier, due to the need of a 10× magnifying lens to examine each kernel. It is therefore necessary to develop a quick, reliable, and objective method to determine if a kernel is infected with *Fusarium* spp. and, consequently, the level of DON content within the kernel (Shahin and Symons, 2011).

There are presently various methods to measure the concentration of DON. These include, but are not limited to: thin-layer chromatography, gas chromatography, mass spectrometry, and enzyme-linked immunosorbent assay (ELISA). Despite the variety of available procedures and their ability to provide an objective measurement, none of the aforementioned methods are suitable for a quick and reliable investigation of DON content (Shahin and Symons, 2011).

2.6.2 Current Applications

The detection of *Fusarium* spp. in grain is an important quality assessment parameter within the agricultural industry. There has been a long-term struggle involved with objectively detecting and identifying *Fusarium* spp. and consequently DON content, in grain commodities. Methods that currently exist for fungal detection include: taxa sequencing, genotypic approaches, and physiological characterization. These approaches provide a valid means of identifying fungus, yet are expensive and time consuming. They are also unable to differentiate between closely related species and provide only a limited amount of data for certain fungal taxa (Santos et al., 2010).

In the past decade, FT-IR spectroscopy has been investigated as a technique to replace conventional means of detecting fungi. More recently, FT-IR spectroscopy has been coupled with superior detection methods, such as FPA and ATR, to provide a rapid, thorough output. FT-IR spectroscopy has been successful in a number of fungal studies including the characterization of *Trametes versicolor* and *Schizophyllum commune* (Naumann et al., 2005), examination of the hyphae of *Aspergillus*, *Neurospora*, and *Rhizopus* (Szeghalmi, 2006), identifying usnic acid in lichen (Liao et al., 2010),

investigating the composition of *Curvularia protuberata* (Isenor et al., 2010), discriminating between different strains of *F. oxysporum* (Salman et al., 2011), and many more. However, before FT-IR spectroscopy could be adequately researched as an agricultural fungal analysis method, sufficient background knowledge of how grain kernels and fungus are detected by FT-IR spectroscopy had to be obtained.

In 1999, Marcott et al. used FT-IR spectroscopy with a 64×64 HgCdTe FPA to determine if structural characteristics of wheat kernels could be modeled. Wheat kernels were cut into 8 μm cross sections and placed onto a BaF₂ window. Spectra were taken in the 900-1800 cm⁻¹ range at 16 cm⁻¹ resolution. FFT and triangular apodization functions were performed on the collected data. It was determined that FT-IR spectroscopy is capable of sampling a small target zone and, when coupled with FPA, a powerful analytical technique arises. This is due to the ability to map the sample with the FPA, as well as understand its chemical composition with the infrared spectrum produced. The work done by Marcott et al. (1999) in this study sets the stage for several other similar experiments, two of which are described below.

Walker et al. (2009) performed a study in which FT-IR spectroscopy was used to determine the structural characteristics of grain endosperm and highlight differences in protein and carbohydrate concentration of corn and barley samples. Five barley varieties and one corn variety were grown, harvested, frozen, and cut into 6 μm sections. Spectra of these sections were taken in the 800-4000 cm⁻¹ range with a 4 cm⁻¹ resolution. Reference measurements, utilizing numerous methods, were taken in order to compare the results obtained by FT-IR spectroscopy. Dry matter content was analyzed using an ultracentrifugal mill, protein content was measured using the Association of Official

Analytical Chemists (AOAC) method, starch content was found using Megazyme Total Starch Assay Kit, and the amount of glucose was measured with a spectrophotometer. The results demonstrated that different types of grain, as well as different varieties of the same grain, have different structural characteristics. This conclusion shows that FT-IR spectroscopy is capable of revealing the structural characteristics of grain, whether it is for different types, varieties, or fungal identification.

Brewer and Wetzel (2010) used FT-IR spectroscopy to determine the phenotypic expression of wheat kernels. They did this by comparing soft glumes, which are desirable for cultivation, with hard glumes, which are considered the wild-type, in order to determine a method in which breeders can use cultivars to thresh both types of wheat at harvest. 9 wild and 11 soft glumes were obtained and phenotypically identified by a wheat breeder. The samples were then stored in airtight conditions, soaked, sliced into 100 μm thick sections, and mounted on BaF₂ windows. Spectra were obtained in the 800-4000 cm^{-1} range with the option for either a single element or a FPA detector. It was found that FT-IR spectroscopy is capable of objectively measuring the toughness of a wheat glume by comparing the lignin content to the cellulose content, which is found from the spectra (Brewer and Wetzel, 2010).

The aforementioned research demonstrates that FT-IR spectroscopy is capable of modeling the biological and structural characteristics of a variety of grain types. With this background, other researchers were able to perform studies in which FT-IR spectroscopy has been used to detect the presence of Fusarium and the content of DON in wheat kernels. Several of these works are outlined below.

Greene et al. (1992) performed a comparison study that used FT-IR spectroscopy combined with photoacoustic spectroscopy (PAS) and diffuse reflectance spectroscopy (DRIFT) in an attempt to differentiate between healthy corn kernels and corn kernels that were artificially contaminated with *F. moniliforme*. The fungal cultures were grown, the mycelia was harvested, washed, and lyophilized. The kernels were inoculated with *F. moniliforme* by adding 5 mL of the fungal solution to a mixture of 50 g of corn kernels and 20 mL of deionized water in a 250 mL Erlenmeyer flask. The kernels were then incubated at 60°C to remove residual water and spectra were collected at 4 cm⁻¹ resolution in the wavelength range of 400-4000 cm⁻¹. The data collected from this method indicated that an uninfected corn kernel could be distinguished from an infected corn kernel. The authors concluded that FT-IR-PAS was the superior technique as it could distinguish between differing corn varieties, did not require any sample preparation, and was able to analyze the whole kernel. However, the authors also note that this methodology is tedious, as it is only able to scan one kernel at a time, and therefore not practical.

Another FT-IR accessory comparison study was done by Abramovic et al. in 2007. This study evaluated the capability of DRIFT spectroscopy to that of ATR coupled with FT-IR spectroscopy in order to detect *Fusarium* spp. on wheat. The wheat samples were artificially inoculated with *F. graminearum*, ground, and scanned in the wavelength range of 650-4000 cm⁻¹, with a resolution of 4 cm⁻¹. A cup holding the ground sample was used for the DRIFT portion of the study and a ZnSe crystal, with twelve internal reflections, was used for the ATR section of the investigation. Reference DON levels were found by subjecting the kernels to high performance liquid chromatography with

diode array detection (HPLC-DAD). Spectra were subject to principal component analysis and the blank and contaminated samples were separated using cluster analysis. The authors found that the ATR method was capable of correctly categorizing 100% of samples, whereas the DRIFT technique was unable to demonstrate clustering of the contaminated and blank samples. Therefore, it was concluded that ATR FT-IR spectroscopy is a superior technique for DON classification of ground wheat kernels.

Santos et al. (2010) compiled a review article detailing the ability of FT-IR spectroscopy to identify and classify fungi. The focus of their review article was to present the current state of FT-IR spectroscopic applications with filamentous fungi and yeasts and draw conclusions based on previous research. The authors provided four references to previous filamentous fungi studies, ranging from the investigation of FT-IR spectroscopy to detect airborne fungi to the detection of fungi in wood, paper, and maize samples. Based on their review, it was concluded that FT-IR spectroscopy provides rapid and reliable data with simple sample preparation when compared to conventional methods. Additionally, FT-IR spectroscopy coupled with ATR, provides the most potential for future filamentous fungi investigations, when compared with other coupling modes.

Kos et al. (2002) reiterated Abramovic and Santos et al.'s conclusion of the superiority of FT-IR spectroscopy coupled with ATR in a study utilizing this technology for the detection of *F. graminearum* on maize kernels. A diamond ATR crystal, capable of producing three internal reflections, was used in conjunction with FT-IR spectroscopy in the wavelength range of 650-4000 cm^{-1} . Maize samples were dried, ground, and then pressed against the ATR crystal for analysis. A pressure applicator was used to ensure the

pressure applied by the ATR crystal was the same for each sample of ground maize. 26 spectra were recorded for each sample. The first derivative of each maize spectrum was taken, followed by principal component analysis. Reference measurements were completed via extraction of the ground maize sample, followed by HPLC-DAD to determine the total fungal biomass in the sample and gas chromatography with electron capture detection (GC-ECD) to determine the sample's DON levels. Upon classifying blank and contaminated samples, it was found that 75% of all samples were correctly classified using the ATR in conjunction with FT-IR spectroscopy. The authors noted that the failure of 100% of the kernels to be correctly classified was due to inhomogeneities that exist within grain kernels themselves.

In 2009, an article was published on the analysis of DON in durum and common wheat. The samples were naturally infected by *F. graminearum* and *F. culmorum* in Italian fields between the years of 2002-2006. A total of 262 samples (143 of durum and 119 of common wheat) were used in the study. The samples were ground with a mill which housed a 500 μm sieve and scanned using FT-NIR spectroscopy (wavelength range of 4,000-10,000 cm^{-1}) with a spectral resolution of 8 cm^{-1} . The samples were also analyzed with HPLC to establish reference DON levels; the authors defined anything with less than 50 $\mu\text{g}/\text{kg}$ as DON-free and anything over that as containing DON. Standard normal variate, first derivative mean centering, and Savitzky-Golay smoothing was performed on the data. Coefficients of determination were calculated for two different parameters: prediction of DON content ($r^2 = 0.71-0.83$) and discrimination between high (>300 $\mu\text{g}/\text{kg}$) and low (<300 $\mu\text{g}/\text{kg}$) DON content ($r^2 = 0.58-0.63$). The authors defined these two coefficients to be “good” and “limited to good”, respectively. Additionally, the

authors found that FT-NIR was capable of separating 69% of contaminated samples from blank ones with a cutoff level of 300µg/kg. The performance of FT-NIR spectroscopy in this application demonstrates that an objective food quality analysis is capable with this technology (De Girolamo et al., 2009).

Utilizing findings from previous research, Kos et al. (2007) describe a sample preparation procedure that can be used in order to determine the DON level in maize using ATR coupled with FT-IR spectroscopy in the mid-infrared region. The authors did this by choosing maize samples that were naturally infected with *F. graminearum*, drying them, grinding them, and then sieving and separating the samples into five different bulk particle sizes: <100µm, 100-250µm, 250-500µm, 500-710µm, 710-1000µm). A ZnSe ATR crystal, capable of three internal reflections, was used and spectral data was collected in the wavelength range of 650-4500 cm⁻¹. Reference data for DON content was collected by employing GC-ECD. The authors compared the sieved samples with non-sieved samples infected with DON. It was concluded that with smaller particle diameters and with a narrowing of particle size distributions, the repeatability of spectral measurements can be improved; particularly, the 100-250 µm range showed the best reproducibility. When solely using the 100-250 µm range, classification of samples was improved to 100%, compared to 79% when using all sieve ranges (Kos et al., 2007).

Lattanzio, et al. (2009) reviewed several analytical methods for the identification of tricothecenes in cereal food products, including FT-IR spectroscopy. It was noted that characteristic absorption bands for DON have been found in the wavelength range below 4000 cm⁻¹, yet when wheat samples were investigated it was determined that the samples needed to be highly contaminated (10 ppm) for DON to be discriminated. Despite this,

the authors concluded that FT-IR spectroscopy has an excellent signal-to-noise ratio with no sample preparation and a high sample throughput. However, the authors warn that this method is highly dependent upon calibration and, subsequently, calibration databases.

The aforementioned studies demonstrate that FT-IR spectroscopy is able to reveal several structural characteristics of wheat, corn, and barley, and therefore has the potential to be a powerful characterization tool for fungal and mycotoxin infected grain commodities. Studies were conducted to determine if FT-IR spectroscopy could decipher between FDK's and healthy kernels using different grain commodities. All of these studies found positive results, indicating that this technology does have the potential for fungal characterization. Additionally, comparison studies of several accessories that can be used with FT-IR spectroscopy when identifying *Fusarium* spp. in different grain commodities were performed by several researchers. The results determined that an ATR accessory is the superior one and holds the most potential for future fungal studies. Using FT-IR spectroscopy to identify the mycotoxin DON proved to be more challenging than simply identifying *Fusarium* spp. It was determined that samples needed to be highly contaminated before DON content could be identified.

The majority of the samples in the highlighted studies were sectioned or ground prior to collecting the data. This may be considered a drawback as it introduces a tedious step into the data collection phase. The reviewed study that kept the samples intact was described as "tedious" as FT-IR spectroscopy was only capable of scanning one sample at a time. The aforementioned studies use both artificially and naturally infected kernels of varying commodities which demonstrate that FT-IR spectroscopy is capable of showing fundamental characteristics regardless of infection method or grain type. FT-IR

spectroscopy was also described as a rapid, accurate technique with good signal to noise ratio. The concern of the reliance on calibration databases was noted; however, the increased use of this technology has led to expanding reference libraries (Santos et al., 2010).

3. METHODOLOGY

3.1 System Design

The microspectroscopic system utilized consisted of a Varian 670-IR FT-IR spectrometer (Agilent Technologies, Melbourne, Australia) in conjunction with a Varian 620-IR FT-IR microscope and a germanium ATR crystal. A grain kernel holder was designed and fabricated to keep the samples in place as the ATR crystal applied pressure during scanning. Varian's Resolutions Pro data acquisition system (Version 5.1.0.829) was used to process the spectral results. All components of the system were housed in a controlled, stable environment at a temperature of 22°C.

3.1.1 Selection of FT-IR

In the agricultural industry today, there are a number of limiting factors involved with determining whether or not a kernel is infected with *Fusarium* spp. and, if so, measuring the concentration of DON, if it exists, within a batch of kernels. Primarily, this grading process, in relation to infection by *Fusarium* spp. and DON content, is tedious and involves the destruction of samples in question. Additionally, identifying *Fusarium* spp. in wheat kernels is subjective and prone to error. The intention of the project is to develop a method capable of eliminating these factors in order to provide a practical means to evaluate if a kernel is infected with *Fusarium* spp. and correlate the DON content to the level of infection.

FT-IR spectroscopy was selected as the measurement mode for a variety of reasons. This technology has the potential to measure the presence of *Fusarium* spp. and the level of DON content within kernels in a non-destructive fashion (Lattanzio et al., 2009). The grain kernel does not need to be modified or prepared in any manner prior to scanning. FT-IR spectroscopy gives rapid, real-time results of the sample that has been scanned and provides a sensitive, objective measurement of the chemical composition of the sample. A microscope was chosen to use in conjunction with the spectroscopic system in order to view the surface of the grain kernels and compare the results to what was seen.

3.1.2 Selection of Measurement Mode

Utilizing FT-IR microspectroscopy, devoid of any accessories, as a measurement mode is an acceptable method for measuring trichothecenes in grain commodities; however, there are a number of accessories that exist which can enhance the data, yielding more accurate results. These additions include, but are not limited to, ATR, DR, specular reflectance, polarizers, and transmission accessories.

Attenuated Total Reflectance is an accessory which allows light to penetrate a few microns into a sample. In order for this to occur, the ATR crystal must be in contact with the sample. Incident light shone into the crystal undergoes total internal reflection and is eventually reflected out of the sample towards a detector (Section 2.3.3). ATR can be used with a wide range of sample types, but is generally recommended for samples that possess high infrared absorption (Kazarian and Chan, 2006).

Diffuse reflectance occurs when the incident beam is reflected off the sample at multiple angles from the sample's surface (Section 2.3.1). Accessories can be purchased that utilize this method during spectroscopic measurement. DR is primarily employed for the measurement of rough surface materials and powders (Milosevic and Berets, 2002).

Specular reflection is a situation in which the incoming beam is reflected off the surface of the sample as a single beam, such that the angle of incidence is equal to the angle of reflection. Specular reflection accessories allow non-destructive collection of data by providing a reflective surface against which the sample can be placed. Due to the mirror-like conditions necessary for specular reflection, the type of samples typically measured with this mode include coatings or thin films (Lellmann et al., 2008).

Transmission occurs when light penetrates the sample and is passed onto a detector on the opposite side of the incident beam (Section 2.3.1). Due to the fact that the incident light must pass through the sample prior to measurement, samples for this application cannot be opaque or completely reflective.

Fig. 3.0 details different technical features available for FT-IR microspectroscopic systems and the type of application they are commonly used for (Varian, 2008). ATR provides the most flexible option of all the methods listed, when considering sample type.

Figure 3.0 Common FT-IR microspectroscopic accessories and corresponding sample types. *Source:* Varian Inc., 2008

Sampling Technique and Common Accessories	Polymers/ Materials	Pharma	Biotech	Chemicals	Academia	Security/ Environment	Forensics
Attenuated Total Reflection (ATR)	●	●	●	●	●	●	●
Diffuse Reflection (DR)	●	●	○	●	●	●	●
Specular Reflectance	●	○	○	●	●	○	○
Grazing Angle Reflectance	●	○	○	●	●	○	○
Microscopy (Single Point)	●	●	●	●	●	●	●
Chemical Imaging (Micro, Macro)	●	●	●	●	●	●	●
ATR Chemical Imaging (Micro, Macro)	●	●	●	●	●	●	●
Fiber Optics (NIR)	○	●	●	●	○	●	○
Fiber Optics (MIR)	○	○	○	●	●	●	○
FT-Raman	●	●	●	●	●	○	●
TGA-FT-IR	●	●	●	●	○	○	●
Polarization Modulation Reflection-Absorption (PM-IRRAS)	●	○	●	○	●	○	○
Photoacoustic Spectroscopy (PAS)	●	○	○	○	●	●	●
Nanosecond Time Resolved Spectroscopy (nsTRS)	○	○	●	○	●	○	○

● Commonly used ● Moderately used ○ Rarely used

Due to the fact that wheat kernels are solid, opaque, non-homogenous samples, and from the findings of current research in the field (Section 2.6.2), ATR was chosen as the mode of measurement. A germanium ATR crystal was used in this study; this allowed for one internal reflection with a penetration of 0.5 to 1 μm into the sample. Germanium crystals are ideal to use with highly absorbing materials (Pike Technologies, 2000), and therefore, are optimal to use when collecting information from whole wheat kernels. Employing the ATR crystal in conjunction with the FT-IR spectral microscope maintains

all of the advantages given by the FT-IR technique on its own, while providing a more sensitive approach by penetrating into the sample.

3.1.3 Design of Kernel Holder

In order to obtain and sustain sufficient contact between the ATR crystal and the whole grain kernel, a means of holding the kernel was required. The intent of the kernel holder was to prevent the kernel from moving when pressure from the ATR crystal was applied. Therefore, the kernel holder needed to remain level under the FT-IR microscope and cover all mirrors and other machinery that are exposed below the slide stage. Due to the fact that the slide stage is equipped to hold a microscope slide, the foundation of the kernel holder design comes from a traditional microscope slide of size 75 mm × 25 mm × 1 mm. The remainder of the stage was measured and the holder was sized to cover and protect the machinery housed below the stage. A slot in the center of the slide portion of the kernel holder was designed in order to support the kernel while it was being scanned. Ten kernels were measured along their largest length and width projections (Table 3.0). It was concluded that ten was a sufficient kernel number for this design as the measurements were becoming redundant at that point. The largest kernel size was used to determine the required slot size. An AutoCad drawing of the kernel holder can be found in Appendix A.

Table 3.0 Measured kernel sizes.

Kernel	Length (inches)	Width (inches)
1	5/16	3/16
2	5/16	2/16
3	1/4	2/16
4	1/4	2/16
5	1/4	2/16
6	1/4	3/16
7	1/4	2/16
8	1/4	2/16
9	1/4	2/16
10	1/4	2/16

3.2 Experimental Design

In order to conduct a reliable study, the type of grain commodity used, division of samples into similar categories, determination of the number of samples to be scanned, and the method by which the samples would be scanned by the FT-IR microspectroscopic unit needed to be established. This section details the decisions and reasoning behind the aforementioned variables.

3.2.1 Samples

Canadian Western Red Spring (CWRS) wheat samples were chosen for this study, as this grain commodity accounts for the largest proportion of grain grown in and exported from Canada (Canadian Wheat Board, 2011). The CWRS samples were obtained from Cargill Ltd. as to approximate as closely as possible the natural conditions of wheat kernels infected with *Fusarium* spp. in agricultural fields. The samples were obtained from nineteen grain elevators across the three Prairie Provinces (Alberta, Saskatchewan, and

Manitoba). The staff at the grain elevators collected the samples from the farmer's truck via a 1L cup, placed the sample in a plastic bag, and sent them via mail to the grain laboratory at Cargill's head office in Winnipeg, Manitoba.

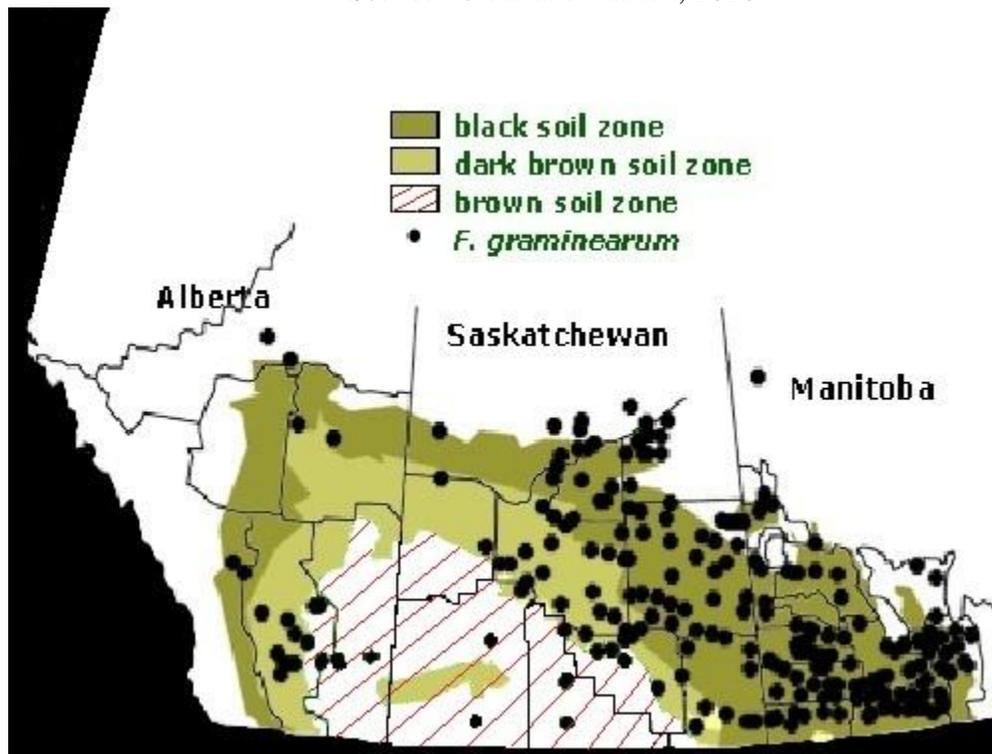
At the grain laboratory, the samples were graded based on the primary grade determinants outlined in the CGC's Official Grain Grading Guide, as shown in Appendix B (Canadian Grain Commission, 2011). This is the standard protocol to grade all Canadian grain. FDK's were identified by a trained, CGC-certified inspector who reduced the sample size received to 50 g and removed the individual kernels that displayed symptoms of being infected with *Fusarium* spp. (see Section 2.6.1). The suspect kernels were then put under a 10× microscope to determine if there was any mould growth on the kernel. The mass of the confirmed FDKs was then recorded and a *Fusarium* damage percentage, based on the 50 g reduced sample size, was obtained as a measure to represent the overall sample.

One hundred and one sample bags containing approximately one kilogram of kernels were received. Inside each sample bag was a smaller bag containing the 50 g reduced sample, and inside each 50 g bag was another bag containing the FDKs found within that sub-sample. The received samples were stored in a refrigerator at a temperature of 0°C. The kernels were kept at this temperature to ensure the fungi could no longer grow and infect the kernel further than it already was when received (Gilbert and Fernando, 2004).

The CGC tracked the distribution of the *F. graminearum* strain of *Fusarium* across the Prairie Provinces between the years of 1994-2008. This study found that FDKs

in the Prairie Provinces primarily reside in three different soil zones, black, dark brown, and brown, as shown in Fig. 3.1. As per the results found by the CGC, the obtained sample bags were divided into the three soil groups, based on the location from which they were received. Separation of the kernels into their respective soil zones subsequently allowed for analysis of kernels that have similar traits with other kernels possessing different characteristics. A summary of the received samples divided into the relevant soil zone can be found in Appendix C.

Figure 3.1 The distribution of *F. Graminearum* across the prairie provinces.
Source: Clear and Patrick, 2010



In the agricultural industry, DON content of a sample of kernels is primarily found by the ELISA procedure. A minimum of 200 g of wheat, containing both uninfected and infected kernels, are ground until 75% of the mass passes through a 20

mesh sieve. ELISA is then performed and the DON content reported back is representative of the entire sample.

Due to the nature of this study, the ELISA procedure was modified in order to determine the DON content in a single infected wheat kernel. The modified procedure consisted of mixing 10 g of ground, uninfected wheat kernels with the single ground infected kernel. 100 mL of distilled water was added to the mixture and shaken for three minutes on a shaker. The mixture was allowed to settle and then was filtered through a cotton ball syringe. After filtration, 100 μ L of the solution was added to the ELISA well and tested. DON contents of each infected kernel were reported back in parts per million (ppm).

3.2.2 Determination of Sample Size

A preliminary experiment was performed in order to determine what, if any, spectral features distinguish a kernel infected with *Fusarium* spp. from one that is uninfected. The pilot study was done with 18 kernels, 9 infected and 9 uninfected. Of the spectra obtained, five functional groups were prominent: CH₂ stretch, lipid carbonyl, protein amide one, cellulose, and the OH-NH stretch.

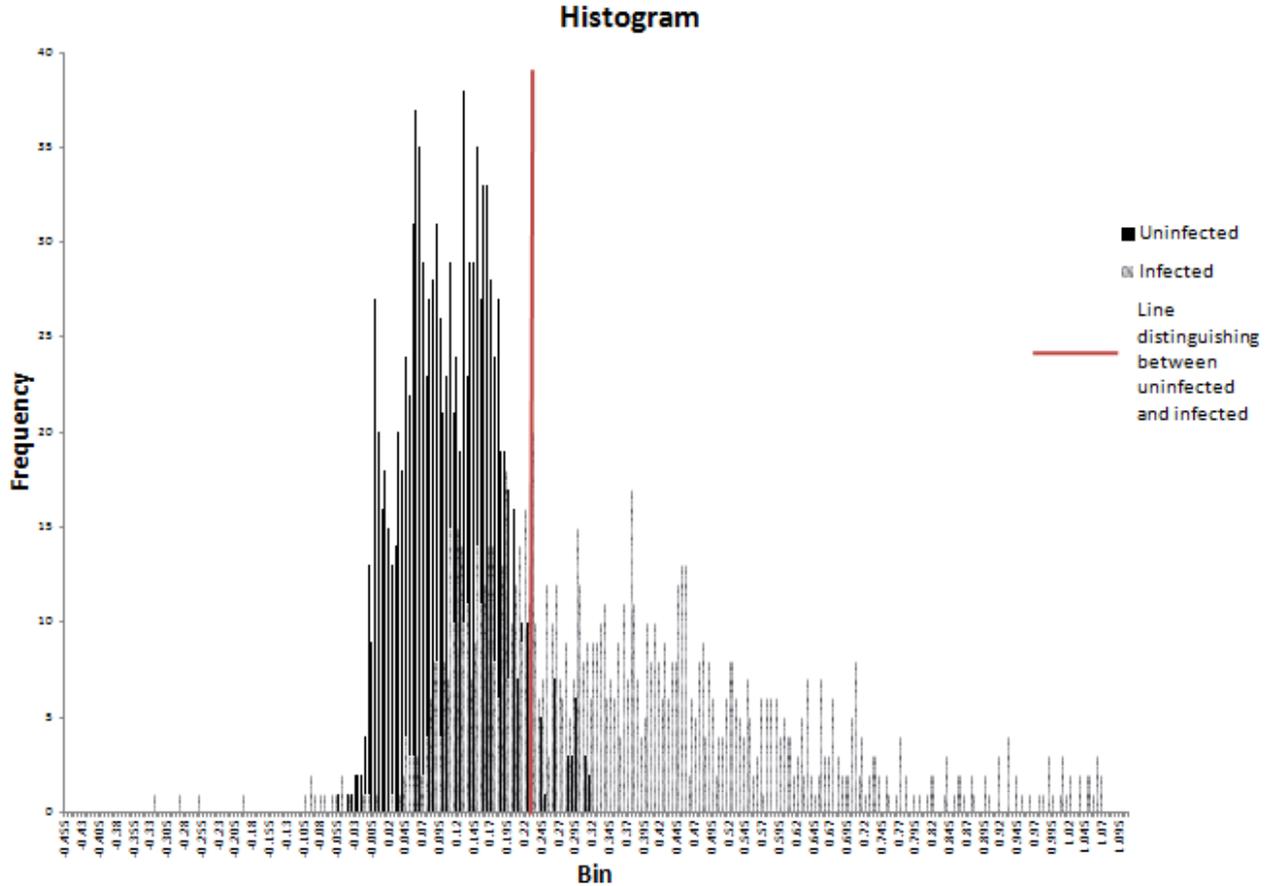
A total of 4096 spectra were collected per wheat kernel and 80 of these were randomly selected as a representative spectral sample of each kernel. Due to the unique shape of whole wheat kernels, some of the spectra were discarded as contact was minimal or absent. Integrations were performed over the full width half maximum (FWHM) of each spectral group for each of the aforementioned functional groups. Results were

summarized, and the mean and standard deviation value for each functional group was calculated from the total number of integrations for all eighteen kernels.

In order to determine what functional group(s), if any, possessed a significant difference between the infected and uninfected kernels, statistical tests were performed. Before the data could be subjected to statistical analysis, it was put through an exploratory phase to determine whether the integrations of each functional group followed a normal distribution. Outlier data was removed via the procedure described by Myers et al. (2010) prior to performing the normality tests. The data was explored by means of histograms, box plots, Q-Q plots, and by performing the univariate procedure in the SAS statistical analysis package (SAS 9.1.3). It was found that the distribution of the integrations for all five functional groups conformed to the normal distribution, which implies that the population data is also likely to follow a normal distribution. Confirming the validity of the normal distribution was necessary to ensure that underlying assumptions involving certain statistical tests and procedures would not be violated.

The histogram of each functional group yielded the type one and type two errors to be used in the statistical analysis. This was done by determining a point on the histogram in which uninfected kernels would be inaccurately classified as infected kernels and vice versa. For example, the point chosen for the protein amide one functional group was an integration value of 0.25 (see Fig. 3.2). Any kernel that possessed an integration value greater than 0.25 for its protein amide one group would be classified as infected, and any kernel with a value below 0.25 would be classified as uninfected.

Figure 3.2 Histogram of protein amide one integrations depicting the area chosen to distinguish the type one and two errors of the study



The type one error was found by counting the number of uninfected kernels that had integration values greater than 0.25 and dividing by the total number of kernels. The type two error was determined by counting the number of infected kernels that had integration values less than 0.25 and dividing by the total number of kernels. Similar procedures were done for the remaining four functional groups.

The data was put through a t-test of independent means. This test was chosen as there are two groups of interest, one with kernels that are infected with *Fusarium* spp. and one control group, which are being analyzed on the same variable of interest, the functional group of choice. In some cases, a modified Welch's t-test was used to account

for unequal variances between the uninfected and infected kernels (Myers et al., 2010). The null hypothesis for each functional group was that the integration value of the uninfected and infected kernels were equal ($\mu_{\text{uninfected}} = \mu_{\text{infected}}$). The alternative hypothesis was that the integration value of the infected kernel was greater than ($\mu_{\text{uninfected}} < \mu_{\text{infected}}$) or less than ($\mu_{\text{uninfected}} > \mu_{\text{infected}}$) that of the uninfected kernel, depending upon which way the distribution of the infected kernels was spread, as seen by the histogram of each functional group. Of the five prominent functional groups, the t-test for the protein amide one and lipid carbonyl functional groups demonstrated that there was a difference between the means of the infected and uninfected kernels, whereas the t-test for the remainder of the functional groups did not find a difference. Due to this result, the protein amide one and lipid carbonyl groups were subject to further statistical analysis to determine what the sample size of the experiment should be.

The sample size was calculated using the power procedure in the SAS statistical analysis package for sample size determination. In order to use this procedure the type one error, the power, the type of test (one-tailed t-test for independent means), the expected range of means and the expected range of standard deviations needed to be input into the program as known variables. To determine what the expected range of means and standard deviations should be, ninety percent confidence intervals (CI) were calculated for the difference between the means and the standard deviations. This was done with the following equation:

$$CI = V \pm (t_{df, \alpha} S_v) \tag{3.1}$$

where V is the variable of interest, $t_{df,\alpha}$ is the critical t-value for the specified degrees of freedom (df) and type one error rate (α), and s_v is the standard error of the variable of interest. The results for the protein amide one and the lipid carbonyl functional group can be found in Fig. 3.3 and 3.4, respectively.

Figure 3.3 SAS power procedure results for protein amide one.

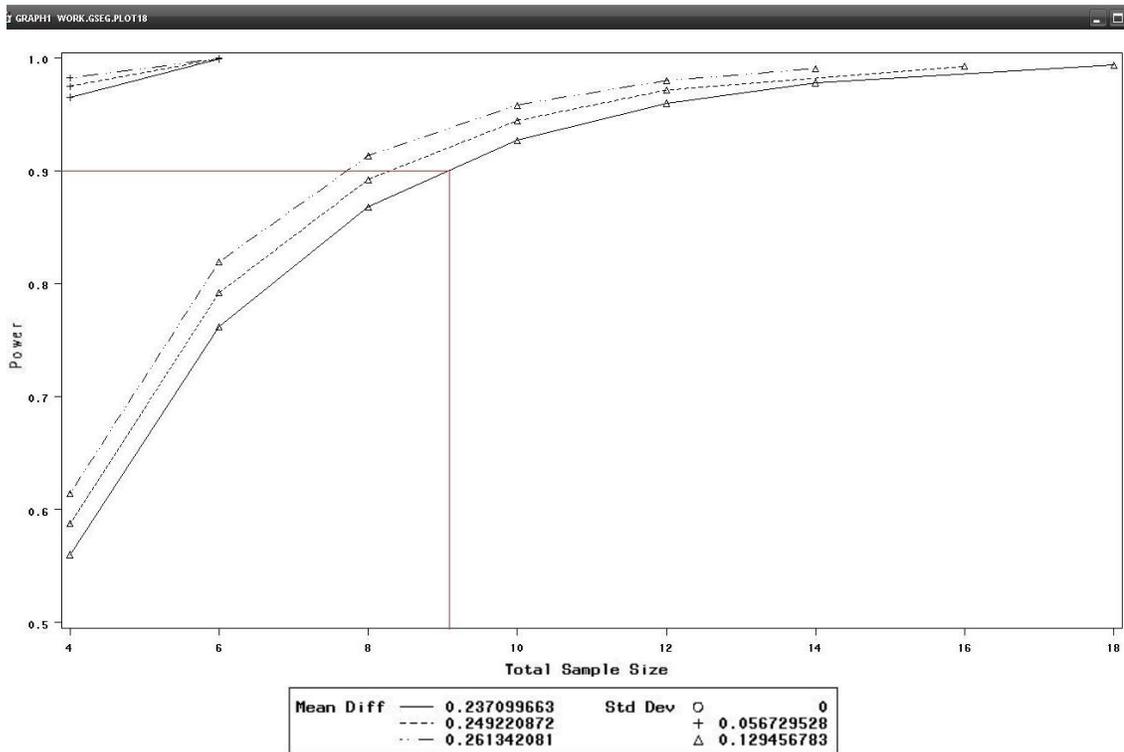
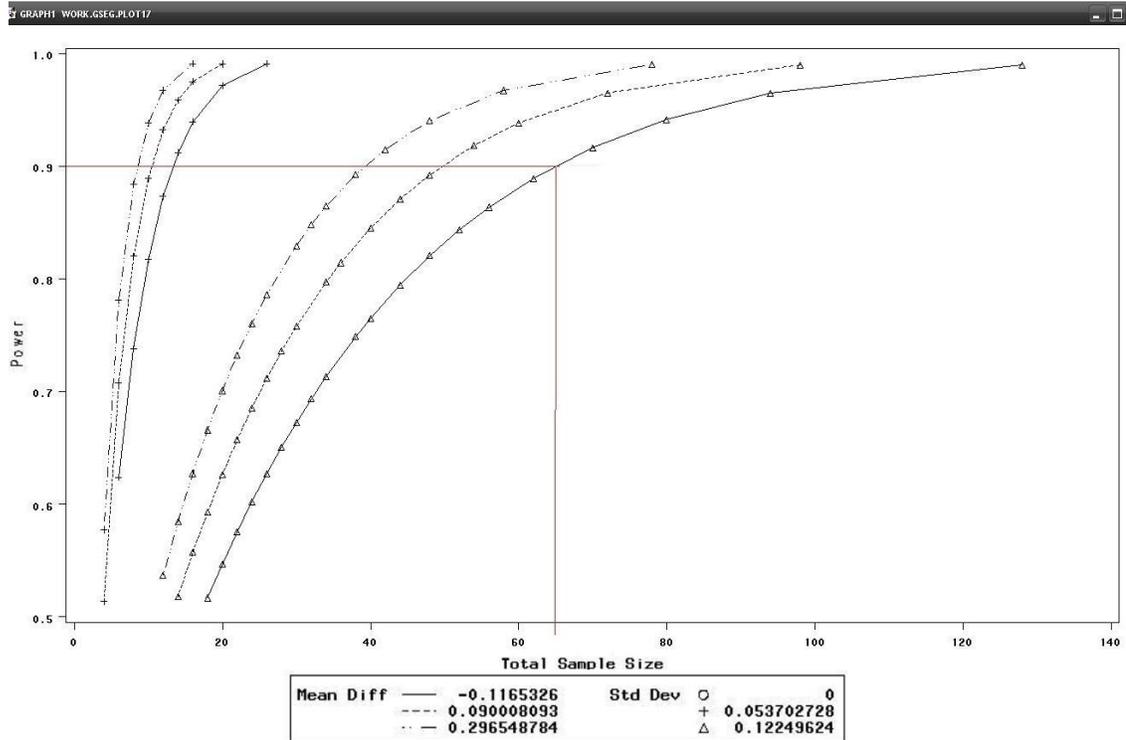


Figure 3.4 SAS power procedure results for lipid carbonyl.



In order to be ninety percent confident that the difference between the uninfected and infected kernels would be accurately detected, the largest sample size found by SAS needed to be used. For the protein amide one functional group, this was approximately 9 uninfected and 9 infected kernels, and for the lipid carbonyl group this was approximately 65 uninfected and 65 infected kernels.

A second calculation was performed using the G*Power 3 statistical analysis tool as recommended by Myers et al. (2010) in order to verify the results from the SAS power procedure. G*Power 3 requires input information about the type one error, the power, and the ratio of the desired size of sample groups to one another (N_2/N_1 , where N_k is sample size of group k), in order to calculate the sample size of the study. Additionally, the effect size (d) of the study is a required input. The effect size is a “scale free index of the

importance of the effect” that is being sought in the experiment (Myers et al., 2010). The effect size can be calculated using Cohen’s equation:

$$d = \frac{(\mu_1 - \mu_2)}{\sigma} \quad (3.2)$$

where $\mu_1 - \mu_2$ is the difference between the uninfected and infected means and σ is the variance. Figure 3.5 and 3.6 show the G*Power 3 result for the protein amide one and the lipid carbonyl groups, respectively.

Figure 3.5 G*Power 3 result for the protein amide one functional group.

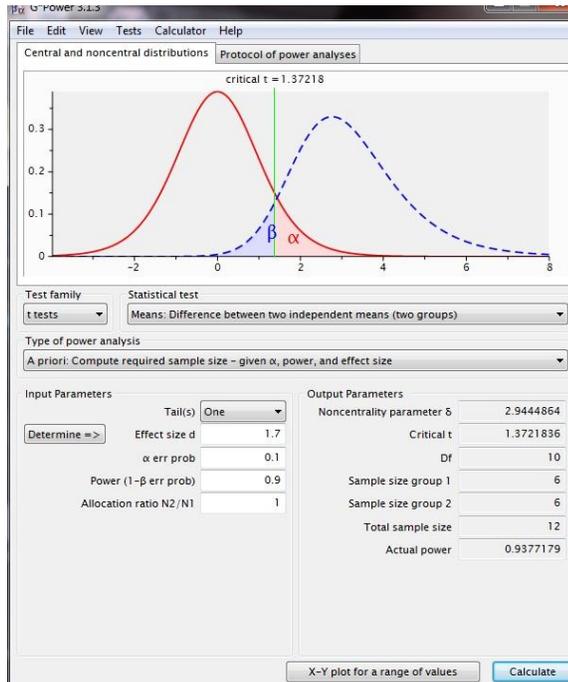
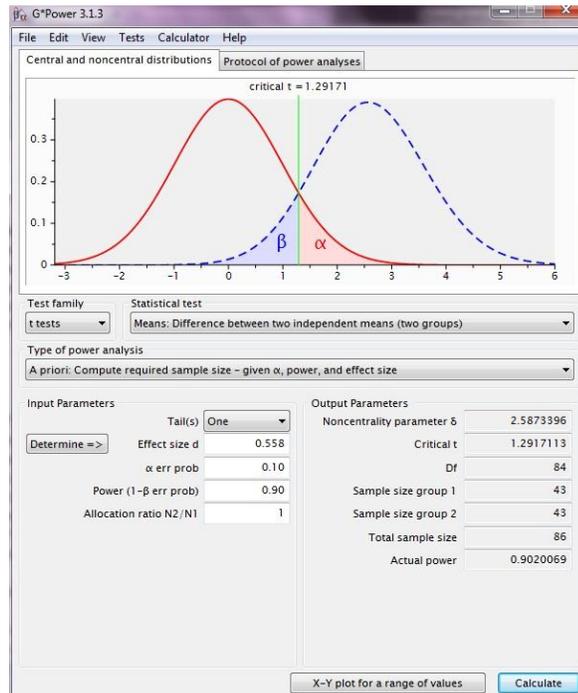


Figure 3.6 G*Power 3 result for the lipid carbonyl functional group.



The G*Power 3 statistical analysis tool yielded a smaller sample size for both the protein amide one and lipid carbonyl functional groups than the SAS power procedure. In order to ensure the sample size accurately reflects the population, the largest sample size (SAS power procedure for lipid carbonyl) was chosen as the one to use for the experiment.

The 130 kernel sample size given by SAS for the lipid carbonyl group did not simply divide into the three soil zones that the samples were separated into (see Section 3.2.1). Therefore, the sample size for this experiment was raised from 130 kernels to 150 kernels in order to ensure that each soil zone had an equal number of representative samples selected from it. A random number generator was used for each soil zone to generate twenty-five random numbers from those assigned. These numbers were matched

up to the corresponding sample bag and one uninfected and one infected kernel was scanned from the respective sample bag. Details of the random numbers generated for each soil zone, as well as the sample bags used in the experiment can be found in Appendix D.

3.2.3 Scanned Samples

Samples were scanned via two methods. The single element (SE) detector was used to determine when contact between the ATR crystal and the grain kernel were made and a FPA of 64×64 size was employed to collect spectral data from the kernel.

Prior to utilizing these methods to collect data from the kernels, a background spectrum of 256 scans was taken for both methods. A spectral resolution of 4 cm^{-1} in the wavelength range of $450\text{-}4400 \text{ cm}^{-1}$ was used for SE, and a resolution of 2 cm^{-1} in the wavelength range and $900\text{-}4000 \text{ cm}^{-1}$ was used for FPA. Spectra collected for the kernels consisted of 128 scans with 4 cm^{-1} resolution for SE and 2 cm^{-1} resolution for FPA, in the same wavelength range that their background spectra were collected from. The discrepancy between the spectral resolution of both methods arises due to the intended use of the method. Since SE was used solely to determine when the ATR crystal made contact with the wheat kernel, the resolution of the method was relaxed.

The samples were placed into the slot of the kernel holder with tweezers. Wheat kernels have a unique external shape; the kernel is elongated, with an oval shaped body and a deep crease on one side of the kernel (Fig. 3.7). The wheat kernels were scanned on the uppermost portion of the body (the endosperm), on the side opposite to the crease. This area was chosen as this is the portion of the kernel where *Fusarium* spp. enter (see

Section 2.6.1) and is the only place on a wheat kernel that gives enough surface area for the ATR crystal to touch down upon without complete interference by the curved surface of the kernel. Therefore, the kernels were placed in the sample holder crease down, to ensure contact with the ATR crystal could be made. Prior to scanning the kernel, the ATR slide was pushed three quarters of the way into its holder with a portion of it open to the atmosphere. It was left in that position to purge for five minutes in order to rid any additional carbon dioxide or water from the region of the ATR slide housing the crystal, prior to taking a measurement. Purging the equipment was necessary to reduce unwanted noise in the final spectra.

Figure 3.7 Top, side, and bottom view of wheat kernel body.



After purging, the ATR crystal was pressed against the uppermost, endosperm region of the kernel. Spectral data was collected; the SE scan taken first, followed by the FPA method. Once scanning was complete, the ATR crystal was removed and its surface cleaned with methanol to eliminate any possible cross-contamination between the samples. The scanned kernel was taken out of the holder, a new one was placed in, the ATR crystal was returned to purge, and the entire process repeated for all of the samples.

3.3 Data Analysis

The data obtained was subject to a spectral and statistical analysis. The spectral analysis was required to gather useful information from the spectra in order to understand the quantity of chemicals within the scanned wheat kernels and any common anomalies that existed in the data. The statistical analysis allowed an exploration of the extracted data to determine whether there are any chemical features within the spectra that indicate the presence of *Fusarium* spp. and, if so, if any of these can be correlated to the amount of DON content found within the kernel.

3.3.1 Spectral Analysis

The Resolutions Pro data acquisition system was used to observe spectra in order to determine the type and quantity of functional groups present in the wheat kernels (Section 3.2.2) and if any common anomalies existed. Rotational-vibrational lines, such as CO₂ and water vapour, were found in some of the spectra; their quantity and presence was not significant enough to warrant any spectral correction techniques. Additionally, spectra with abnormal baselines were found where the ATR pressure was at its maximum throughout all of the collected data (Section 4.1.1). In some of the spectra, CH₂ stretching was found to possess high absorption values. It was therefore decided that this functional group would need to be further investigated.

The G*Power 3 software, as recommended by Meyers et al. (2010), was used to determine the number of spectra used per kernel. The specified test was an analysis of variance (ANOVA) F-test. The inputs required for the F-test were the effect size, type one error, power, and the number of groups in the study. The effect size, type one error,

and power variables were determined from the pilot study. The number of groups input into the software was three, due to the three soil zones the samples came from. G*Power3 output a result of 207 spectra for the three groups which gave a value of 69 spectra each. This was rounded to 70 for ease of collection.

From previous research (Section 2.6.1), and from the results of the pilot study, the lipid carbonyl, protein amide one, and cellulose conglomerate functional groups were further investigated to determine if they indicate the presence of *Fusarium* spp. within wheat kernels. Additionally, the asymmetric and symmetric CH₂ stretch functional groups were examined to establish whether these groups indicate the presence of *Fusarium* spp. Spectra were first viewed using the Resolutions Pro data acquisition system in order to determine if any significant differences between the uninfected and infected kernels could be seen by the unaided eye. It was ultimately determined that statistical analysis was needed to investigate the data further (Section 4.1.1). The spectra were then loaded into Matlab software (Mathworks Inc., Natwick, MA) Version 7.11.0.584 in order to perform integrations over the FWHM of the appropriate functional groups. Integrations were calculated in order to determine the quantity of the functional group within the kernel. Table 3.1 lists the functional group, the left and right baseline, and the left and right edges used for each FWHM calculation. In addition to integrating over the FWHM of the curve, ratios between the different spectral parameters were performed, via the integration values. These ratios gave an indication of how the functional groups differ from one another, based on a single reference peak (cellulose conglomerate functional group).

Table 3.1 Functional group with left and right baseline of FWHM.

Functional Group	Left Baseline (cm⁻¹)	Left Edge (cm⁻¹)	Right Edge (cm⁻¹)	Right Baseline (cm⁻¹)
Lipid Carbonyl	1811	1745	1716	1692
Protein Amide One	1691	1666	1613	1610
Cellulose	1192	1111	970	920
CH ₂ Stretch (Symmetrical)	2871	2858	2846	2831
CH ₂ Stretch (Asymmetrical)	2991	2951	2897	2885

For each functional group, MATLAB randomly selected 70 spectra from each kernel. These spectra were inspected and if any displayed poor contact or irregular baselines they were discarded. Additional random spectra were generated in Matlab until 70 spectra that demonstrated good kernel contact and baselines were obtained for every spectral parameter. These randomly generated pixels were compiled and a total of 350 spectra were obtained for each functional group, for every kernel, and carried through to the analysis phase.

3.3.2 Statistical Analysis

Prior to performing any statistical analysis, the data was examined in an exploratory phase to determine if it conformed to the normal distribution. This phase ultimately dictated what statistical tests can be used, and if any modifications needed to be made to the traditional tests. The exploratory phase was conducted by examining histograms, box plots, stem-and-leaf plots, Q-Q plots, normal probability plots, and examining the variance, skewness, and kurtosis values of each group. The data collected was divided into three soil zones and analyzed for the five different functional groups (see Section 3.3.1). In order to analyze a test of this nature, a one-way ANOVA F-test was performed.

The F-test was used to determine if a difference between the means of the infected and uninfected samples in each soil zone, based on the five different spectral parameters tested, exists. The F-test was performed for both the integration and ratio values. If a significant result was found from the ANOVA F-test for either method, further evaluation was performed via contrasts.

Each significant parameter found was contrasted. The contrast was performed for all of the uninfected and infected kernels of each soil zone (ie. lipid carbonyl uninfected vs. lipid carbonyl infected) in order to narrow down the results of the ANOVA F-test and determine which functional group, if any, indicates the presence of *Fusarium* spp. Depending on the results, further contrasts between each soil zone were performed to determine if the location the kernel was grown in had an effect on the ability of the spectra to indicate the existence of the fungus within the kernel.

Based on the results from the contrasts, a correlation between the functional group(s) that indicated the presence of *Fusarium* spp. and the level of DON content found in the kernel was carried out. This was done by employing several different methods (see Section 4.2). From these results, it can be determined whether any relationship between the level of DON content and the functional group of interest exists.

4. RESULTS AND DISCUSSION

4.1 Exploration of Data

An experiment was conducted to determine whether *Fusarium* spp. could be detected on whole wheat kernels via FT-IR microspectroscopy, coupled with an ATR accessory. A total of 4096 spectra were collected for 150 wheat kernels. The array of spectra obtained all displayed similar characteristics, which dictated the functional groups of interest for further investigation (Section 3.2.2). The spectra were put through an image exploratory phase which revealed some key instrumental and spectral characteristics and confirmed the need for a statistical analysis. A total of 70 random spectra for each functional group were selected and compiled, resulting in a total of 350 spectra for every kernel. The area under each spectral curve of interest was integrated at FWHM to obtain a quantitative value of the absorption of that particular spectral peak. Additionally, ratios of these integrated values were calculated against the cellulose conglomerate functional group apex, as this peak demonstrated consistency throughout the spectra in terms of its relative absorbance to the other peaks, as well as the height and width of the peak itself. The concentration of DON within the kernels was obtained via the procedure described in Section 3.2.1.

4.1.1 Image Exploration

A total of 5.25×10^4 spectral images were obtained for the uninfected and infected kernel data set. These spectral images were explored in order to determine if any common trends or anomalies existed within the data set. The majority of the spectra obtained exhibited a

common curve, typical of the chemical constituents found within a wheat kernel. As previously mentioned, the functional groups of interest in this study are the cellulose conglomerate, the CH₂ asymmetric and symmetric stretch, protein amide one, and lipid carbonyl.

Several features were found in the spectral images that were common among all of the collected data. One of these features was the presence of an ATR pressure point that was located in the center of each frequency slice, across many pixels (Fig. E.0). Spectral data within this ATR pressure point exhibited abnormal baselines and thus were discarded if they had been randomly selected (Appendix E). It is likely that this pressure point was created due to the curved nature of the wheat kernel surface. Only good quality spectra that exhibited a good SNR and level baseline, outside of this pressure point region, were chosen for further analysis.

A second prominent feature was the appearance of striped patterns in the images created from the frequency slices for each of the functional groups of interest. This striated appearance conforms to the curved, grooved surface of a wheat kernel and was an expected result of ATR contact in this study. Appendix F details typical spectra, including zoomed in views of the functional groups of interest, their frequency slices, and the wheat kernel surface as seen by a 4× microscope, for uninfected and infected kernels in the black, brown, and dark brown soil zones.

As can be seen from the spectra in Appendix F, it is difficult to discern with the unaided eye whether significant differences between the uninfected and infected wheat kernels exist based on the five selected functional groups. It was therefore necessary to

turn to a statistical analysis to further investigate the data. This was done by calculating the FWHM of the five functional groups for every 350 spectra collected for each wheat kernel in order to get a value of absorption that could be used in the statistical examination.

4.1.2 Statistical Exploration

Data exploration must occur before starting analysis to determine if the desired inferential statistical tests are applicable and what modifications, if any, need to be made. The objective of the exploration phase is to confirm the normality and homogeneity of variance properties of the data. In order to explore the data, the distribution of observations should be graphed in numerous ways and several key descriptor statistics, such as the mean and standard deviation, must be calculated (Myers et al., 2010).

Data detailing the spectral parameter of concern, the soil zone of the kernel, the infection level of the kernel (uninfected or infected), the integrated value of the spectral parameter, and the ratio of the functional groups was uploaded into SAS for the exploratory phase. This phase was carried out graphically via histograms, stem-and-leaf plots, boxplots, Q-Q plots, and normal probability plots. Additionally, descriptor statistics were calculated, namely the mean, standard deviation, skewness, and kurtosis values. Furthermore, the Shapiro-Wilk test of normality was performed.

The graphical exploration phase revealed that all of the data trended towards the normal distribution, with slight skewness and kurtosis. This was verified by calculation of several statistical descriptor values for the integration (Table 4.0) and the ratio (Table 4.1) values. The skewness and kurtosis values of each spectral group, for both the

integration and ratio values, were found to be small and a modified test was not required based on these two factors. The mean of both the integration and ratio values of each spectral parameter showed common trends between the infected and uninfected kernels, and further investigation into this descriptor was not required. The variance of the infected and uninfected kernels in each spectral parameter and soil group, for both the integration and ratio values, are not identical. The significance of this variance must be tested for in order to determine whether a modification to the one-way ANOVA F-test will need to be made. This will be done by performing Levene's test for homogeneity of variance. If the assumption of homogeneity of variance is violated, Welch's ANOVA F-Test will be executed to compensate for the violation.

The Shapiro-Wilk test of normality rejected the null hypothesis that the data came from a normally distributed population for all of the data in the experiment. However, the Shaipro-Wilk results can be misleading in instances where there is a large data set, "as small departures from normality may yield significant results" (Myers et al., 2010). It is likely that the small skewness and kurtosis values found for the data are responsible for the Shapiro-Wilk results; nevertheless further investigation into the violation of normality was performed.

Table 4.0 Statistical descriptor results of the integrated values from the exploratory phase.

Spectral parameter	Soil zone	Fusarium infection	Mean	Standard deviation	Skewness	Kurtosis
Cellulose Conglomerate	Black	Uninfected	24.03	5.85	-0.46	-0.36
Cellulose Conglomerate	Black	Infected	18.57	7.80	-0.24	-0.53
Cellulose Conglomerate	Brown	Uninfected	22.86	6.30	-0.56	-0.21
Cellulose Conglomerate	Brown	Infected	17.48	7.10	-0.44	-0.25
Cellulose Conglomerate	Dark Brown	Uninfected	21.77	7.83	-0.73	-0.05
Cellulose Conglomerate	Dark Brown	Infected	19.77	6.61	-0.36	-0.28
CH ₂ Asymmetric Stretch	Black	Uninfected	1.31	0.42	-0.29	-0.08
CH ₂ Asymmetric Stretch	Black	Infected	1.08	0.62	0.13	-0.64
CH ₂ Asymmetric Stretch	Brown	Uninfected	1.38	0.48	-0.33	-0.12
CH ₂ Asymmetric Stretch	Brown	Infected	1.11	0.62	-0.23	-0.39
CH ₂ Asymmetric Stretch	Dark Brown	Uninfected	1.25	0.51	-0.17	-0.51
CH ₂ Asymmetric Stretch	Dark Brown	Infected	1.05	0.58	0.08	-0.58
CH ₂ Symmetric Stretch	Black	Uninfected	0.20	0.07	0.10	-0.23
CH ₂ Symmetric Stretch	Black	Infected	0.16	0.10	0.28	-0.44
CH ₂ Symmetric Stretch	Brown	Uninfected	0.21	0.08	-0.14	-0.10
CH ₂ Symmetric Stretch	Brown	Infected	0.19	0.11	0.10	-0.53
CH ₂ Symmetric Stretch	Dark Brown	Uninfected	0.18	0.09	-0.05	-0.27
CH ₂ Symmetric Stretch	Dark Brown	Infected	0.15	0.09	0.33	-0.44
Lipid Carbonyl	Black	Uninfected	0.49	0.24	-0.06	-0.18
Lipid Carbonyl	Black	Infected	0.38	0.29	0.34	-0.36
Lipid Carbonyl	Brown	Uninfected	0.54	0.31	0.18	-0.41
Lipid Carbonyl	Brown	Infected	0.41	0.29	0.36	-0.21
Lipid Carbonyl	Dark Brown	Uninfected	0.50	0.30	0.25	-0.25
Lipid Carbonyl	Dark Brown	Infected	0.42	0.39	0.93	0.92
Protein Amide 1	Black	Uninfected	0.67	0.26	0.08	0.05
Protein Amide 1	Black	Infected	0.55	0.38	0.24	0.18
Protein Amide 1	Brown	Uninfected	0.47	0.39	-0.22	-0.01
Protein Amide 1	Brown	Infected	0.45	0.53	-0.27	0.04
Protein Amide 1	Dark Brown	Uninfected	0.48	0.33	-0.11	-0.15
Protein Amide 1	Dark Brown	Infected	0.65	0.44	0.21	-0.19

Table 4.1 Statistical descriptor results of the ratio values from the exploratory phase.

Spectral parameter	Soil zone	Fusarium infection	Mean	Standard deviation	Skewness	Kurtosis
CH ₂ Asymmetric Stretch	Black	Uninfected	1.31	0.42	-0.29	-0.80
CH ₂ Asymmetric Stretch	Black	Infected	1.08	0.62	0.13	-0.64
CH ₂ Asymmetric Stretch	Brown	Uninfected	1.38	0.48	-0.33	-0.12
CH ₂ Asymmetric Stretch	Brown	Infected	1.11	0.62	-0.23	-0.39
CH ₂ Asymmetric Stretch	Dark Brown	Uninfected	1.25	0.51	-0.51	-0.17
CH ₂ Asymmetric Stretch	Dark Brown	Infected	1.05	0.58	0.08	-0.58
CH ₂ Symmetric Stretch	Black	Uninfected	0.20	0.07	0.10	-0.23
CH ₂ Symmetric Stretch	Black	Infected	0.16	0.10	0.28	-0.44
CH ₂ Symmetric Stretch	Brown	Uninfected	0.21	0.08	-0.14	-0.10
CH ₂ Symmetric Stretch	Brown	Infected	0.19	0.11	0.10	-0.53
CH ₂ Symmetric Stretch	Dark Brown	Uninfected	0.18	0.09	-0.05	-0.27
CH ₂ Symmetric Stretch	Dark Brown	Infected	0.15	0.10	0.33	-0.44
Lipid Carbonyl	Black	Uninfected	0.49	0.24	-0.06	-0.18
Lipid Carbonyl	Black	Infected	0.38	0.29	0.34	-0.36
Lipid Carbonyl	Brown	Uninfected	0.54	0.31	0.18	-0.41
Lipid Carbonyl	Brown	Infected	0.41	0.29	0.36	-0.21
Lipid Carbonyl	Dark Brown	Uninfected	0.49	0.30	0.25	-0.25
Lipid Carbonyl	Dark Brown	Infected	0.39	0.34	0.55	-0.17
Protein Amide 1	Black	Uninfected	0.67	0.26	0.08	0.05
Protein Amide 1	Black	Infected	0.55	0.38	0.24	0.18
Protein Amide 1	Brown	Uninfected	0.47	0.39	-0.22	-0.01
Protein Amide 1	Brown	Infected	0.45	0.53	-0.27	0.04
Protein Amide 1	Dark Brown	Uninfected	0.48	0.33	-0.11	-0.15
Protein Amide 1	Dark Brown	Infected	0.65	0.44	0.21	-0.19

Meyers et al. (2010) recommend three procedures for dealing with violations of normality. These are transformations, trimming outliers, and tests based on ranks. Three transformations were applied to the data in an attempt to achieve normality: square root, logarithmic, and inverse transformations. When exploring the data after these transformations had been performed, it was found that the transformations caused the data to further deviate from normality. Outliers were trimmed via the method recommended by Meyers et al. (2010). This resulted in a marginal shift towards normality, which was still rejected by the Shapiro-Wilk test. Despite this rejection, the trimmed data was carried out through the remainder of the investigation as the data was better aligned with the normal distribution.

The final procedure to deal with violations from normality is to perform a rank-transformed test. However, the traditional F test is relatively unaffected from normality deviations, unless the departure is quite extreme or the sample size is small (Myers et al., 2010). Despite this, a rank-transformed test was carried out, in addition to the traditional F-test. This was to ensure that if the detected violations from normality were a concern, they would be compensated for through the ranked test. The Kruskal-Wallis test is one of the most common rank-transformed F tests and was chosen to apply to the data.

4.2 Detection of *Fusarium* spp. in Whole Wheat Kernels

Data was partitioned into the five functional groups of interest and the soil zone of the kernel, the infection level of the kernel (uninfected or infected), the integrated value of the respective spectral parameter, and the ratio values were uploaded into SAS for significance calculations. Since outlier data was removed from the set, the PROC GLM

command was used to perform the one-way ANOVA F-test, as this test compensates for unbalanced data. Levene's test was also performed, along with Welch's ANOVA F-test to deal with heterogeneity of variance issues as they arose. The PROC NPAR1WAY procedure was used to perform the Kruskal-Wallis test. The type one error rate was set at 0.1 for all tests performed, as dictated by the pilot study. The model for the ANOVA calculations specified the spectral integration value as the dependent variable, and soil, infection, and the interaction between soil and infection as the independent variables. Table 4.2 outlines the p-value results from the integration values used in the study. The p-value is the probability that, assuming the null hypothesis is true (ie. that the observations are equal, in this case the means of the infected and uninfected kernels in each soil zone are equal), the test statistic is at least as extreme as the data obtained. The null hypothesis is rejected if the p-value is less than the type one error rate of the study (Meyers et al., 2010), in this case, less than 0.10. For example, a p-value of <0.0001 means there is less than a 0.01% probability that the parameters tested are actually the same.

Table 4.2 P-value results of several statistical tests for the integration values.

Spectral parameter	Source	ANOVA	Levene's *	Welch's*	Kruskal- Wallis*
Cellulose Conglomerate	Model	<0.0001	N/A	N/A	N/A
Cellulose Conglomerate	Soil Zone	<0.0001	0.0016	<0.0001	<0.0001
Cellulose Conglomerate	Infection	<0.0001	<0.0001	<0.0001	<0.0001
Cellulose Conglomerate	Soil Zone*Infection	<0.0001	<0.0001	<0.0001	N/A
CH ₂ Asymmetric Stretch	Model	<0.0001	N/A	N/A	N/A
CH ₂ Asymmetric Stretch	Soil Zone	<0.001	<0.0001	<0.001	<0.0001
CH ₂ Asymmetric Stretch	Infection	<0.0001	<0.0001	<0.0001	<0.0001
CH ₂ Asymmetric Stretch	Soil Zone*Infection	<0.0001	<0.0001	<0.0001	N/A
CH ₂ Symmetric Stretch	Model	<0.0001	N/A	N/A	N/A
CH ₂ Symmetric Stretch	Soil Zone	<0.0001	<0.0001	<0.0001	<0.0001
CH ₂ Symmetric Stretch	Infection	<0.0001	<0.0001	<0.0001	<0.0001
CH ₂ Symmetric Stretch	Soil Zone*Infection	<0.0001	<0.0001	<0.0001	N/A
Lipid Carbonyl	Model	<0.0001	N/A	N/A	N/A
Lipid Carbonyl	Soil Zone	<0.0001	<0.0001	<0.0001	<0.0001
Lipid Carbonyl	Infection	<0.0001	<0.0001	<0.0001	<0.0001
Lipid Carbonyl	Soil Zone*Infection	0.0002	<0.0001	<0.0001	N/A
Protein Amide 1	Model	<0.0001	N/A	N/A	N/A
Protein Amide 1	Soil Zone	<0.0001	<0.0001	<0.0001	<0.0001
Protein Amide 1	Infection	0.0112	<0.0001	0.0022	0.1677
Protein Amide 1	Soil Zone*Infection	<0.0001	<0.0001	<0.0001	N/A

*Levene's, Welch's, and Kruskal-Wallis tests are able to evaluate one parameter at a time

To summarize, the overall model ANOVA F-test for all spectral parameters was found to be significant, indicating that there is a difference between the means of the infected and uninfected kernels. Additionally, each parameter in the model was also found to be significant. The tests of homogeneity of variance failed across all of the functional groups, yet Welch's ANOVA F-test did not change the significance found in the traditional ANOVA F-test. This is likely due to the large number of data points used in the experiment, and the sensitivity of Levene's test. The Kruskal-Wallis test further confirmed the results found from the other two tests among all of the functional groups. The only exception to this was the infection parameter in the protein amide one

functional group, which did not find the means of the uninfected and infected kernels to be significantly different. From these results, it was determined that contrasts needed to be performed to narrow down the significant results. The contrasts were also performed in SAS 9.1.3 by using the CONTRAST command and choosing the desired parameter to be contrasted. The result of the contrast function is a p-value which determines whether the means of the parameters contrasted are equal or not, based on their value and the type one error rate of the study.

The total numbers of infected and uninfected kernels for each functional group were contrasted. Additionally, the total number of infected and uninfected kernels in each separate soil zone were contrasted, for each functional group, to determine if the type of soil had an effect on the detection of Fusarium in the kernel. It was found that all contrasts were significant at the $p < 0.0001$ level, except the total number of infected and uninfected kernels for the protein amide one functional group. With a p-value of 0.0112, it is still considered significant at the type one error rate of this study. From these results, it was determined that further investigation into how the integration value of each functional group is related to DON content was needed (Section 4.2).

The ratio values were also subject to the aforementioned tests that were performed on the integration values. Table 4.3 summarizes the results found for the ratio values.

Table 4.3 P-value results of several statistical tests for the ratio values.

Spectral parameter	Source	ANOVA	Levene's *	Welch's*	Kruskal- Wallis*
CH ₂ Asymmetric Stretch	Model	<0.0001	N/A	N/A	N/A
CH ₂ Asymmetric Stretch	Soil Zone	<0.0001	<0.0001	<0.0001	<0.0001
CH ₂ Asymmetric Stretch	Infection	<0.0001	<0.0001	<0.0001	<0.0001
CH ₂ Asymmetric Stretch	Soil Zone*Infection	<0.0001	<0.0001	<0.0001	N/A
CH ₂ Symmetric Stretch	Model	<0.0001	N/A	N/A	N/A
CH ₂ Symmetric Stretch	Soil Zone	0.0015	0.1212	0.0022	<0.0001
CH ₂ Symmetric Stretch	Infection	0.0416	0.1496	0.0414	0.0002
CH ₂ Symmetric Stretch	Soil Zone*Infection	0.0086	0.0452	<0.0001	N/A
Lipid Carbonyl	Model	0.0010	N/A	N/A	N/A
Lipid Carbonyl	Soil Zone	0.0015	0.1364	0.0011	<0.0001
Lipid Carbonyl	Infection	0.2602	0.6871	0.2608	<0.0001
Lipid Carbonyl	Soil Zone*Infection	0.0446	0.0038	<0.0001	N/A
Protein Amide 1	Model	0.0099	N/A	N/A	N/A
Protein Amide 1	Soil Zone	0.5106	0.3335	0.2915	<0.0001
Protein Amide 1	Infection	0.0040	0.1211	0.0036	<0.0001
Protein Amide 1	Soil Zone*Infection	0.0714	0.3703	<0.0001	N/A

*Levenes, Welchs, and Kruskal-Wallis tests are able to evaluate one parameter at a time

The ratio values demonstrated that all model ANOVA F-tests were found to be significant. However, the infection parameter of lipid carbonyl and the soil parameter of protein amide one were not found to be significantly different. From Levene's test, it was observed that homogeneity of variance held for all of the functional groups except the CH₂ asymmetric stretch group, and the interaction between the soil and infection level for the CH₂ symmetric stretch and lipid carbonyl functional groups. Welch's ANOVA F-test did not change the outcome of the significance of any parameters tested. The Kruskal-Wallis test results were different for the two nonsignificant parameters and found that the means of these two parameters differed in terms of the uninfected and infected kernels.

The significance of these results and the contradiction between the two ANOVA F-tests and the Kruskal-Wallis test prompted the need for contrasts to be performed

between the ratio values. The same parameters that were contrasted for the integration values are used for the ratio values and the results are shown in Table 4.4.

Table 4.4 P-value results of contrast tests performed for ratio values.

Spectral parameter	Contrast performed	P-value
CH ₂ Asymmetric Stretch	Infected vs. Uninfected	<0.0001
CH ₂ Asymmetric Stretch	Black Soil Infected vs. Black Soil Uninfected	<0.0001
CH ₂ Asymmetric Stretch	Brown Soil Infected vs. Brown Soil Uninfected	<0.0001
CH ₂ Asymmetric Stretch	Dark Brown Soil Infected vs. Dark Brown Soil Uninfected	<0.0001
CH ₂ Symmetric Stretch	Infected vs. Uninfected	0.0416
CH ₂ Symmetric Stretch	Black Soil Infected vs. Black Soil Uninfected	0.9285
CH ₂ Symmetric Stretch	Brown Soil Infected vs. Brown Soil Uninfected	0.0002
CH ₂ Symmetric Stretch	Dark Brown Soil Infected vs. Dark Brown Soil Uninfected	0.0052
Lipid Carbonyl	Infected vs. Uninfected	0.2602
Lipid Carbonyl	Black Soil Infected vs. Black Soil Uninfected	0.7434
Lipid Carbonyl	Brown Soil Infected vs. Brown Soil Uninfected	0.0071
Lipid Carbonyl	Dark Brown Soil Infected vs. Dark Brown Soil Uninfected	0.6867
Protein Amide One	Infected vs. Uninfected	0.0040
Protein Amide One	Black Soil Infected vs. Black Soil Uninfected	0.9281
Protein Amide One	Brown Soil Infected vs. Brown Soil Uninfected	0.0502
Protein Amide One	Dark Brown Soil Infected vs. Dark Brown Soil Uninfected	0.0016

The contrasts show that the ratio values for the total numbers of the uninfected and infected kernels are significantly different from one another for each functional group, except for the lipid carbonyl group. However, when contrasts between the ratio values of the kernels in each soil zone were performed, it was found that the infected and uninfected kernels of the brown soil zone were significantly different for the lipid carbonyl group, meaning a kernel could be distinguished as uninfected or infected based on this parameter, but only within the brown soil zone. Additionally, the ratio values of the uninfected and infected kernels in the black soil zone for the CH₂ symmetric stretch,

lipid carbonyl, and protein amide one functional groups were not significantly different from one another, which indicates that a kernel could not be differentiated as uninfected or infected based on these parameters. Despite the fact that this result was found for the lipid carbonyl and several functional groups of black soil zone, further investigation into how the ratio values of the above functional groups may correlate with DON content was performed (Section 4.2). The results for both the integration and the ratio values demonstrate a significance when observing the interaction between the soil zone the kernel was located in and the level of infection (uninfected or infected) of the kernel. This implies that the soil zone the wheat kernel was grown in affects the ability to distinguish between an uninfected and infected kernel. This significance is explored in the next section.

4.3 Effects of Canadian Soil Zones

Western Canada is divided into different soil zones based on climate and vegetative factors that ultimately determine the amount of organic material found in the topsoil of a particular geography (Government of Manitoba, 2010). The CGC found that the occurrence of *Fusarium* spp. primarily resided in three of these soil zones: black, brown, and dark brown (Section 3.2.1). The kernels used in this study were partitioned into these three soil zones for further investigation. Due to the results found in Section 4.1.2, it was necessary to explore the effects of each soil zone for the ability to correctly classify infected and uninfected kernels and determine how this classification relates to the soil zone.

The kernels within the three soil zones were investigated for their ability to distinguish between the means of the infected and uninfected kernels for each spectral parameter (Section 4.1.2). Additionally, the infected and uninfected kernels of each soil zone were contrasted with the infected and uninfected kernels of the other two zones in order to investigate the significant interaction effect found between soil and infection. The p-value results of the contrasts are summarized in Table 4.5 and 4.6 for the integration and ratio values, respectively.

Table 4.5 P-value results of soil specific contrast results for integration values.

Contrast performed	Cellulose Conglomerate	CH ₂ Asymmetric Stretch	CH ₂ Symmetric Stretch	Lipid Carbonyl	Protein Amide One
Black Infected vs. Other Infected	0.5749	0.9542	<0.0001	<0.0001	0.6676
Brown Infected vs. Other Infected	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
Dark Brown Infected vs. Other Infected	<0.0001	<0.0001	<0.0001	0.0011	<0.0001
Black Uninfected vs. Other Uninfected	<0.0001	0.6278	0.2005	<0.0001	<0.0001
Brown Uninfected vs. Other Uninfected	0.6875	<0.0001	<0.0001	<0.0001	<0.0001
Dark Brown Uninfected vs. Other Uninfected	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

Table 4.6 P-value results of soil specific contrast results for ratio values.

Contrast performed	CH ₂ Asymmetric Stretch	CH ₂ Symmetric Stretch	Lipid Carbonyl	Protein Amide One
Black Infected vs. Other Infected	0.9542	0.0052	0.0159	0.0200
Brown Infected vs. Other Infected	<0.0001	<0.0001	<0.0001	0.5613
Dark Brown Infected vs. Other Infected	<0.0001	0.0609	0.0617	0.0780
Black Uninfected vs. Other Uninfected	0.6278	0.5501	0.4766	0.4847
Brown Uninfected vs. Other Uninfected	<0.0001	0.7550	0.4479	0.9417
Dark Brown Uninfected vs. Other Uninfected	<0.0001	0.7716	0.9645	0.4363

The results of Table 4.5 and 4.6 show that, for the majority of situations, there is a significant difference between the infected kernels among the different soil zones. The black soil zone is the exception, indicating that when it comes to some functional groups the infected kernels in this soil zone do not differ from the infected kernels in the other two soil zones. Additionally, the ratio value for the protein amide one functional group of infected kernels in the brown soil zone does not differ from the ratio values for the protein amide one functional group of the infected kernels in the other two soil zones.

When comparing integration values the uninfected kernels are typically different among the soil zones. The uninfected kernels of black soil zone is the exception, in terms of CH₂ asymmetric and symmetric stretches and the uninfected kernels of the brown soil zone in terms of cellulose content. However, the results for the ratio values show that the uninfected kernels do not differ between functional groups among the different soil zones, the CH₂ asymmetric functional group of the uninfected kernels for the brown and

dark brown soil zones being the exception. These results show the need to keep the data from the different soil zones partitioned for the correlation analysis.

Research was done in order to determine why the type of soil zone a kernel was grown in has an effect on the chemical composition of the infected kernels. The climate and vegetative factors of the soil zone were considered. However, this research study did not yield any significant results and it is recommended that a laboratory study which investigates this further be performed (Section 6).

As can be noted throughout this section, the integration and ratio values differ and in some cases give different significance results. This is likely due to an issue of ATR contact with the kernel surface. The integration values take into account the degree of contact between the ATR crystal and the wheat kernel surface, whereas the ratio values do not. Therefore, it can be argued that the ratio values provide more accurate information. Despite this, in order to provide a complete review and understanding of how the functional groups may be correlated with DON content, both integration and ratio values are used.

4.4 Prediction of DON Content from Infection by *Fusarium* spp.

In order to determine whether DON content can be predicted by the spectra obtained from whole wheat kernels, a correlation analysis was performed. This was done to the integration values of five functional groups, namely cellulose conglomerate, CH₂ asymmetric stretch, CH₂ symmetric stretch, lipid carbonyl, and protein amide one. Additionally, the potential for a correlation was explored with the ratio values found for the following four functional groups: CH₂ asymmetric stretch, CH₂ symmetric stretch,

lipid carbonyl (brown soil zone kernels only), and protein amide one. The correlation analysis was performed for each of the aforementioned integration and ratio values for each different soil zone.

A scatter plot was created which compared the amount of DON content and the integration or ratio value of each spectral parameter for each soil zone. The results of this found the highest correlation values, among all of the soil zones, to be: $R^2 = 0.1987$ and $R^2 = 0.1952$, respectively, for the integration and ratio values. These values demonstrate that no immediate correlation exists.

A second approach was taken by utilizing a technique known as a “moving average”. This method is useful in trying to determine underlying relationships by removing complexity from the graph (Meyers et al., 2010). In order to do this, the dependent variable was divided into a set number of bins, and the DON content was averaged over these bins. The bins used for each functional group are detailed in Table 4.7 for both the integration and ratio values. These were determined by noting the spread between the highest and lowest values and dividing that based on providing approximately six bins per graph in order to give enough data points to investigate a correlation.

Table 4.7 Bin sizes for moving average technique.

Functional group	Integration bin size	Ratio bin size
Cellulose Conglomerate	6	N/A
CH ₂ Asymmetric Stretch	0.50	0.05
CH ₂ Symmetric Stretch	0.15	0.03
Lipid Carbonyl	0.30	0.02
Protein Amide One	0.30	0.03

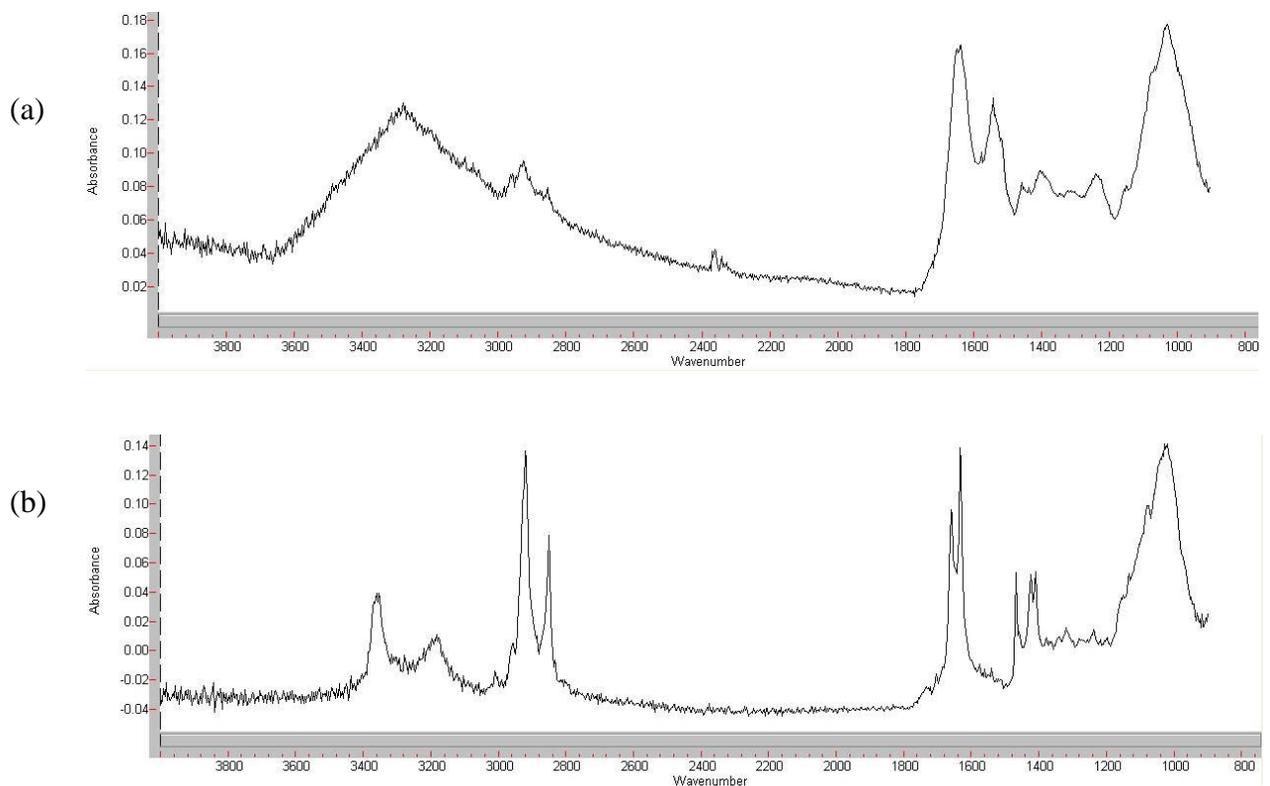
The highest R^2 value results found for this method were $R^2 = 0.7998$ and $R^2 = 0.7877$ for the integration and ratio values, respectively. This method helped to clean up the data and achieve high coefficients of determination. However, since there are 350 independent integration and ratio values for every dependent DON content value, a proper correlation could not be conceived as there are multiple ways a correlation line could run through the data. In order to eliminate this, the integration and ratio values were averaged over each kernel. This averaging was carried out in the remainder of the analysis

A correlation was performed using the average values for each kernel and DON content. The highest R^2 results from all of the data for this method are $R^2 = 0.4302$ and $R^2 = 0.6543$ for the integration and ratio values, respectively. Again, the results do not show a satisfactory correlation between the integration or ratio value and DON content.

It was found from Section 4.1.2 that the means of the uninfected kernels for the integration and ratio values are different from the means of the infected kernels. Another method that attempted to correlate the integration and ratio values to DON content was performed that utilized this fact. The difference between the infected and uninfected kernels was found by taking the mean of the integration or ratio value for the uninfected kernels and subtracting each integration or ratio value of the infected kernels from that average. These differences were then averaged for each kernel. The results of this method yielded an R^2 value of $R^2 = 0.4312$ and $R^2 = 0.3555$ for the integration and ratio values, respectively. Again, these values demonstrate there is no correlation between the integration or the ratio values and DON content.

It is possible that there may be several uninfected spectra in the infected spectral data set, due to the non-homogenous characteristic of the wheat kernel surface. A method that attempted to eliminate any uninfected spectra that may have been included in the data set was then employed in order to determine if this would provide a correlation to DON content. The first step in this approach involved finding spectra of an infected kernel that were significantly different from that of an uninfected kernel. Figure 4.1 shows two such spectra that were found.

Figure 4.0 Spectrum of an infected kernel that is significantly different than the spectra of uninfected kernels showing (a) markedly increased protein amide one content and (b) markedly increased lipid content.



Several spectra were manually chosen that demonstrated the characteristics found in Figure 4.1(a) and (b). The FWHM integration values of these spectra were then taken, ratio values were calculated, and the mean, standard deviation, and 90% confidence

intervals were calculated (Table 4.8). Any integration or ratio value that fell within this confidence interval was chosen for further analysis. This method eliminated many of the spectra, and in some instances, it eliminated all of the spectra and thus no further analysis could be performed. The average integration and ratio values were then taken over each kernel and coefficients of determination were calculated (highest R^2 values of $R^2 = 0.3885$ and $R^2 = 0.6296$ for the integration and ratio values, respectively for (a) and highest R^2 values of $R^2 = 0.3408$ and $R^2 = 0.5207$, respectively for the integration and ratio values for (b)). These values show that no satisfactory correlation to DON content exists. It should be noted that data sets in which only two data points fell within the desired interval were omitted. This is due to the fact that a satisfactory correlation cannot be considered when the data is so reduced.

Table 4.8 Mean, standard deviation, and confidence limits of spectra significantly different than the spectrum of an uninfected kernel.

Spectrum used	Functional group	Integration values				Ratio values				
		Mean	Standard Deviation	Upper Confidence Limit	Lower Confidence Limit	Mean	Standard Deviation	Upper Confidence Limit	Lower Confidence Limit	
Increased Protein Amide One Content	Cellulose	9.8695	0.5891	10.3519	9.3870	N/A	N/A	N/A	N/A	
	CH ₂ Asymmetric Stretch	0.4091	0.0505	0.5262	0.2919	0.4133	0.0027	0.04758	0.0351	
	CH ₂ Symmetric Stretch	0.0166	0.0094	0.0336	-0.0005	0.0017	0.0009	0.0033	6.2538x10 ⁻⁷	
	Lipid Carbonyl Protein Amide One	-0.4387	0.0367	-0.3903	-0.4871	-0.0445	0.0037	-0.0397	-0.0493	
Increased Lipid Content		0.0166	0.0094	0.0242	0.0089	0.0016	0.0009	0.0024	0.0009	
	Cellulose	14.375	3.0478	16.8717	11.8798	N/A	N/A	N/A	N/A	
	CH ₂ Asymmetric Stretch	8	3.2939	0.4017	4.2553	2.3624	0.2372	0.0595	0.3751	0.0992
	CH ₂ Symmetric Stretch	0.6770	0.0943	0.8486	0.5054	0.0489	0.0136	0.0737	0.0242	
	Lipid Carbonyl Protein Amide One	-0.1007	0.1026	0.0346	-0.2361	-0.0079	0.0076	0.0022	-0.0179	
		3.5612	0.3256	3.8279	3.2945	0.2553	0.0546	0.2999	0.2106	

Another approach, similar to the previous one, was explored. This method compared the means of the integration and ratio values of the infected and uninfected kernels. If the mean for a given functional group decreased in infected kernels when compared to uninfected kernels, it was assumed that all integration or ratio values less than the upper confidence limit calculated in the previous method were infected. If the mean for the functional group increased for infected kernels, it was assumed that all integration or ratio values higher than the lower confidence limit were infected. Table 4.9 summarizes the integration and ratio means for the infected and uninfected kernels as well as which confidence limit was used.

Table 4.9 Mean values and confidence limits used – elimination of clean pixels using upper or lower confidence limit.

Soil Zone	Functional group	Integration values			Ratio values		
		Uninfected mean	Infected mean	Confidence Limit	Uninfected mean	Infected mean	Confidence Limit
Black	Cellulose Conglomerate	23.8308	18.5734	Upper	N/A	N/A	N/A
	CH ₂ Asymmetric Stretch	1.3349	1.1073	Upper	0.0644	0.0667	Lower
	CH ₂ Symmetric Stretch	0.2102	0.1675	Upper	0.0100	0.0095	Upper
	Lipid Carbonyl	0.4925	0.3900	Upper	0.0235	0.0256	Lower
	Protein Amide One	0.6492	0.5940	Upper	0.0304	0.0332	Lower
Brown	Cellulose Conglomerate	22.5533	17.2234	Upper	N/A	N/A	N/A
	CH ₂ Asymmetric Stretch	1.3661	1.1321	Upper	0.0774	0.1790	Lower
	CH ₂ Symmetric Stretch	0.2141	0.1915	Upper	0.0123	0.0328	Lower
	Lipid Carbonyl	0.5465	0.4231	Upper	0.03075	.05637	Lower
	Protein Amide One	0.4531	0.4539	Lower	0.0220	0.0649	Lower
Dark Brown	Cellulose Conglomerate	21.3026	19.5167	Upper	N/A	N/A	N/A
	CH ₂ Asymmetric Stretch	1.24305	1.0559	Upper	0.08077	0.08166	Lower
	CH ₂ Symmetric Stretch	0.1852	0.1545	Upper	0.01251	0.01180	Upper
	Lipid Carbonyl	0.5203	0.4175	Upper	0.02899	0.02549	Upper
	Protein Amide One	0.4890	0.6762	Lower	0.01719	0.08596	Lower

The results of this method yielded the highest R^2 values of $R^2 = 0.1855$ and $R^2 = 0.3621$ for the integration and ratio values, respectively for (a) and the highest R^2 values of $R^2 = 0.2017$ and $R^2 = 0.3622$, respectively for the integration and ratio values for (b). Similar to the previous method, this one failed to find a satisfactory correlation and data sets that contained only two kernels within the desired interval were omitted.

A final approach was taken that arose from a study completed by Abromovic et al. (2007). This study utilized FT-IR spectroscopy with ATR and found that a correlation between DON content and the ratio of the absorbance bands of 1709 cm^{-1} to 1743 cm^{-1} could be obtained. The two absorbance values were obtained, the ratio was calculated, and the results were averaged over each kernel. This method yielded the worst correlation to DON content with the highest R^2 value being $R^2 = 0.0827$.

This thesis mimicked the condition of grain kernels as found in the agricultural industry. As a result, there are several sources of uncertainty involved with this project that may have accounted for the inability to establish a correlation between the significant functional groups and DON content. These include, but are not limited to, unknown condition of the plant during growth and harvest. The weather conditions, as well as the time of day and year the plant was initially seeded and finally harvested in. The degree of infection by *Fusarium* spp. within the wheat kernels was unknown, as well as the length of time the kernel had been infected prior to instilling conditions that halted fungal growth. Finally, how the kernels were handled prior to receiving them was unknown.

5. CONCLUSIONS

A FT-IR spectrometer and microscope was used, in conjunction with ATR, to scan whole wheat kernels in an attempt to detect the presence of the *Fusarium* fungus. It was found that the CH₂ asymmetric and symmetric stretches, lipid carbonyl, protein amide one, and cellulose conglomerate functional groups were prominent in every spectrum obtained. The spectra were investigated and it was found that some common trends existed within the data including the presence of an ATR pressure point and the appearance of striated frequency slices for each of the investigated spectral parameters. From the image investigation, it was found that a statistical analysis was necessary to discern differences between the uninfected and infected kernels. This was done by integrating at FWHM under each curve and ratioing that value against the cellulose conglomerate functional group for the remaining four parameters. Both the integration and ratio values were subject to an ANOVA F-test. It was found from the integration values that the uninfected kernels and FDKs could be distinguished from one another based on all of the aforementioned functional groups. The ratio values were used to discriminate between the two types of kernels based on the CH₂ asymmetric stretch, CH₂ symmetric stretch, and protein amide one functional groups.

The ANOVA test found the effect of the interaction between the level of infection of the kernel (uninfected or infected) and the type of soil zone the kernel was grown in (black, brown, or dark brown) to be significant for all functional groups. Contrasts were performed to compare the uninfected and infected kernels of one soil zone to those kernels of the other two soil zones. The contrasts of the integration values revealed that the infected kernels in the black soil zone did not differ from the other infected kernels in

terms of cellulose ($p = 0.5749$), CH₂ asymmetric stretch ($p = 0.9542$), and protein amide one ($p = 0.6676$) content. Additionally, the uninfected kernels of the black soil zone did not differ from the other uninfected kernels in terms of CH₂ asymmetric ($p = 0.6278$) and symmetric stretch ($p = 0.2005$). Furthermore, the uninfected kernels of the brown soil zone were the same as those of the other soil zones in terms of cellulose ($p = 0.6875$) content. The rest of the contrasts revealed significant differences ($p < 0.1000$) for the other functional groups in the other soil zones.

The contrasts of the ratio values found most of the infected kernels in one soil zone to be significantly different ($p < 0.1000$) than the infected kernels of the other two soil zones for all functional groups. The black and brown infected kernels broke this majority in terms of the CH₂ asymmetric stretch ($p = 0.9542$) and protein amide one ($p = 0.5613$) functional groups, respectively. The ratio values found the majority of the uninfected kernels in one soil zone to be the same as the uninfected kernels of the other soil zones for all functional groups. The exception to this was the brown and dark brown soil zones in terms of the CH₂ asymmetric stretch ($p < 0.0001$) functional group. The results for both the integration and ratio values demonstrate a significant difference in the ability to detect *Fusarium* spp. based on the type of soil zone the kernel was grown in.

Several attempts were made to correlate the integration and ratio values to the level of DON content in the kernel. These attempts failed to find a correlation for all functional groups, in all soil zones. This may be due to the non-homogeneous and complex nature of the grain kernel surface, the fact that not all FDKs produce the DON toxin, or that the functional groups that detect *Fusarium* spp. are not a sound predictor for the level of DON content in wheat kernels. Due to the inability to develop a correlation

model for DON content, there was no subsequent test performed to confirm the validity of such an equation.

6. RECOMMENDATIONS

The ability to distinguish between uninfected kernels and FDKs with FT-IR microspectroscopy coupled with an ATR accessory is desirable as it provides an accurate, objective measurement of the kernel in question. It was demonstrated that this discrimination could be made on the basis of the CH₂ asymmetric stretch, CH₂ symmetric stretch, lipid carbonyl, protein amide one, and cellulose functional groups. However, several modifications to the instrument need to be made, prior to utilizing in an industrial setting:

1. Stability: currently, FT-IR spectrometers must be housed in a controlled environment. This is not suitable for the type of environment that is typical of country and terminal grain elevators. The dust, noise, and mechanical and temperature fluctuations that are characteristic of the agricultural industry would affect the accuracy and reliability of the spectrometer; and
2. Speed: although FT-IR spectrometers are capable of providing a rapid scan of the sample in question, this speed needs to be increased prior to industrial use. In this study, it took approximately 20 minutes to collect all 4096 spectra from one kernel at 2 cm⁻¹ resolution. This is much too slow for industrial use. Additionally, more than one kernel would need to be scanned at a time to provide an accurate representation of a batch of kernels.

It was determined that the soil zone in which the kernel was grown has an effect on the ability to classify kernels as uninfected or infected. It would be desirable to further investigate this conclusion in order to understand how soil type affects the functional group parameters that indicate infection by *Fusarium* spp. Additionally, understanding

how to eliminate this effect in order to provide a measurement that is independent of the location of the grain kernel would be beneficial. Alternatively, a software input could be created that would allow the user to specify the location the kernel came from and return a result as to whether or not the kernel is infected with *Fusarium* spp.

A satisfactory correlation between one or more functional groups that indicate infection by *Fusarium* spp. and the level of DON content within the kernel was not achieved. There are several reasons for this, including the non-homogeneous and complex nature of the wheat kernel, the possibility of toxins other than DON being present, and that the functional groups that indicate the presence of *Fusarium* spp. may not be good predictors of DON content. Several possible directions to overcome these challenges in future experiments are given:

1. Minimize the uncertainties that arose in this experiment by artificially inoculating the grain kernels and keeping their surrounding conditions known throughout the study (lab-based study).
2. Attempt to measure past the outer bran layer of the grain kernel. The *Fusarium* fungus does destroy the outer layers and the cell walls of the host kernel; however this destruction is related to the level of infection of the kernel and it is possible that, in this experiment, some of this layer is still intact. Providing a measurement technique that always measures past the bran layer would reduce variation in the spectra obtained. This could be done by employing a different ATR crystal.
3. Develop a software program that can objectively analyze the spectra and distinguish between spectra that demonstrate satisfactory ATR contact and

those that demonstrate poor or no ATR contact. The software program could specify a desired SNR value, as well as baseline tolerances.

4. Develop a software program that can objectively analyze the spectra and distinguish between uninfected and infected spectra of a kernel, ultimately eliminating any uninfected spectra that were collected on a FDK.
5. Measure values of DON content within a kernel by employing a method other than ELISA, such as thin-layer or gas chromatography, to provide more accurate concentrations of DON content.

7. REFERENCES

- Abdi, H. 2003. Partial least squares (pls) regression. *Encyclopaedia of Social Sciences and Research Methods* 1-7.
- Abramovic, B., Jajic, I., Abramovic, B., Cosic, J., and V. Juric. 2007. Detection of deoxynivalenol in wheat by fourier transform infrared spectroscopy. *Acta Chimica Slovenica* 54: 859-867.
- Agriculture and Agri-Food Canada. 2011. An overview of the Canadian agriculture and Agri-Food system 2011. 11279E. Ottawa, Ontario: Agriculture and Agri-Food Canada.
- Åmand, L.E., and C.J.Tullin. 1999. The theory behind ftir analysis. 1-15. Department of Energy Conservation: Chalmers University of Technology, Göteborg, Sweden.
- Anderson, P.E., D.A. Mahle, T.E. Doom, N. V. Reo, N.J. DelRaso, and M.L. Raymer. 2011. Dynamic adaptive binning: an improved quantitative technique for nmr spectroscopic data. *Metabolomics* 7(2): 179-190.
- Azzouz, T., A. Puigdomenech, M. Aragay, and R. Tauler. 2003. Comparison between different data pre-treatment methods in the analysis of forage samples using near-infrared diffuse reflectance spectroscopy and partial least-squares multivariate calibration method. *Analytica Chimica Acta* 484(1):121-134.
- Bambery, K.R., B.R. Wood, and D. McNaughton. 2011. Resonant Mie scattering (RMieS) correction applied to ftir images of biological tissue samples. *Analyst* 137: 126-132.
- Barnes, R. B. and R. C. Gore. 1949. Infrared spectroscopy. *Analytical Chemistry* 21(1): 7-12.

- Barnes, R.J., M.S. Dhanoa, and S.J. Lister. 1993. Correction to the description of standard normal variate (snv) and de-trend (dt) transformations in practical spectroscopy with applications in the food and beverage analysis – 2nd edition. *NIR News* 5(3): 185-186.
- Bassan, P., A. Kohler, H. Martens, J. Lee, E. Jackson, N. Lockyer, P. Dumas, M. Brown, N. Clarke and P. Garner. 2010a. RMieS-EMSC correction for infrared spectra of biological cells: extension using full Mie theory and gpu computing. *Journal of Biophotonics* 3(8-9): 609-620.
- Bassan, P., A. Kohler, H. Martens, J. Lee, H.J. Byrne, P. Dumas, E. Gazi, M. Brown, N. Clarke and P. Gardner. 2010b. Resonant Mie scattering (RMieS) correction of infrared spectra from highly scattering biological samples. *Analyst* 135(2): 268-277.
- Benoudjit, N., F. Melgani and H. Bouzgou. 2009. Multiple regression systems for spectrophotometric data analysis. *Chemometrics and Intelligent Laboratory Systems* 95: 144-149.
- Betchel, D.B., L.A. Kaleikau, R.L. Gaines, and L.M. Seitz. 1985. The effects of fusarium graminearum infection on wheat kernels. *Cereal Chemistry* 62(3): 191-197.
- Biselli, S., L. Hartig, H. Wegner, and C. Hummet. 2005. Analysis of fusarium toxins using lc-ms-ms: application to various food and feed matrices. *Spectroscopy* 20(2): 20-26.
- Biswal, P.C. 2008. *Ordinary Differential Equations*. New Delhi, India: Prentice Hall of India Pvt.Ltd.

- Bokobza, L. 1998. Near infrared spectroscopy. *Journal of Near Infrared Spectroscopy* 6(1): 3-17.
- Bosco, G. L. 2010. Meeting report – James L. Waters symposium 2009 on near-infrared spectroscopy. *Trends in Analytical Chemistry* 29(3): 197-208.
- Boyacioglu, D. and N.S. Hettiarachchy. 1995. Changes in some biochemical components of wheat grain that was infected with fusarium graminearum. *Journal of Cereal Science* 21(1): 57-62.
- Bradner, J.E., and T.Y. Susskind. 1966. Infrared spectroscopy: theory. In *The Electromagnetic Spectrum*. 155-163. Litton Instructional Materials, Inc.
- Brewer, L.R. and D.L. Wetzel. 2010. Phenotypic expression in wheat revealed by using ft-ir microscopy. *Vibrational Spectroscopy* 52(1): 93-96.
- Canadian Grain Commission. 2009. Canadian grain exports: crop year 2009. 1701-9931. Winnipeg, Manitoba: Canadian Grain Commission.
- Canadian Grain Commission. 2011. Wheat – Chapter 4: Official Grain Grading Guide. <http://www.grainscanada.gc.ca/oggg-gocg/04/oggg-gocg-4f-eng.htm>. (2011/02/11).
- Canadian Wheat Board. 2008. The wheat quality control system in Canada. http://www.cwb.ca/public/en/library/research/popups/wheat_Quality_standards.js p. (2011/02/11).
- Canadian Wheat Board. 2011a. About us: frequently asked questions. <http://www.cwb.ca/public/en/about/FAQ/>. (2011/02/11).
- Canadian Wheat Board. 2011b. CWB year-end news conference: crop year 2010-2011. Winnipeg, Manitoba: Canadian Wheat Board.

- Carter, R.L. 1998. *Molecular Symmetry and Group Theory*. New York, NY: John Wiley & Sons, Inc.
- Clear, R. and S. Patrick. 2010. Fusarium head blight in western Canada.
<http://www.grainscanada.gc.ca/str-rst/fusarium/fhbwc-foc-eng.htm>. (2011/05/23).
- De Girolamo, A., V. Lippolis, E. Nordkvist, and A. Visconti. 2009. Rapid and non-invasive analysis of deoxynivalenol in durum and common wheat by fourier-transform near infrared (ft-nir) spectroscopy. *Food Additives and Contaminants* 26(6): 907-917.
- De Wolf E.D., L.V. Madden, and P.E. Lipps. 2003. Risk assessment models for wheat fusarium head blight epidemics based on within-season weather data.
Phytopathology 93(4): 428-435.
- Desjardins, A.E., T.M. Hohn and S.P. McCormick. 1993. Tricothecene biosynthesis in fusarium species: chemistry, genetics, and significance. *Microbiological Reviews* 57(3): 595-604.
- Dexter, J.E., R.M. Clear, and K.R. Preston. 1996. Fusarium head blight: effect on the milling and baking quality of some canadian wheats. *Cereal Chemistry* 73(6): 695-701.
- Doymaz, F., A. Palazoglu, and J.A. Romagnoli. 2003. Orthogonal nonlinear partial least squares regression. *Industrial and Chemical Engineering Research* 42(23):5836-5849.
- Duhamel, P. and M. Vetterli. 1999. Fast fourier transforms: a tutorial review and a state of the art. In *The Digital Signal Processing Handbook*, ed. Madisetti, V.K. and D. B. Williams, 259-299. Boca Raton, FL: CRC Press Taylor & Francis Group.

- Duygu D., T. Baykal, I. Acikgöz, and K. Yildiz. 2009. Fourier transform infrared (FT-IR) spectroscopy for biological studies. *Gazi University Journal of Science* 22(3): 117-121.
- Fletcher, J. 2008. Multiple linear regression. *British Medical Journal* 338(3): 167.
- Gable, K. 2000. FTIR Spectroscopy.
<http://www.chem.orst.edu/courses/ch361-464/ch362/irinstrs.htm> (2010/11/30).
- Geladi P. and B. Kowalski. 1986. Partial least squares regression: a tutorial. *Analytical Chemistry Acta* 185(1): 1-17.
- Gilbert, J. and W.G.D. Fernado. 2004. Epidemiology and biological control of gibberella zea/fusarium graminearum. *Canadian Journal of Plant Pathology* 26: 464-472.
- Global Security. 2010. Infrared Tutorial.
<http://www.globalsecurity.org/space/library/report/1998/sbirs-brochure/part02.htm> (2010/01/03).
- Goswami, R.S. and H.C. Kistler. 2004. Heading for disaster: fusarium graminearum on cereal crops. *Molecular Plant Pathology* 5(6): 515-525.
- Government of Manitoba. 2010. Soil Management Guide: Understanding the Soil Landscapes of Manitoba.
<http://www.gov.mb.ca/agriculture/soilwater/soilmgmt/fsm01s01.html#variation>. (2011/10/24).
- Greene, R.V., S.H. Gordon, M.A. Jackson and G.A. Bennett. 1992. Detection of fungal contamination in corn: potential of FTIR-PAS and-DRS. *Journal of Agricultural Food Chemistry* 40, 1144e1149.

- Griffiths, P. R. and J.A. de Haseth. 2007. *Fourier Transform Infrared Spectroscopy*, 2nd edition. Hoboken, NJ: John Wiley & Sons, Inc.
- Guo, H., K.J. Marfurt, and J. Liu. 2009. Principal component spectral analysis. *Geophysics*. 74(4): 35-43.
- Isenor, M., S.G.W. Kaminskyj, R.J. Rodriguez, R.S. Redman and K.M. Gough. 2010. Characterization of mannitol in *Curvularia protuberata* hyphae by ftir and raman spectromicroscopy. *Analyst* 135: 3249-324.
- Kang, Z. and H. Buchenauer. 2000. Ultrastructural and cytochemical studies on cellulose, xylan and pectin degradation in wheat spikes infected by *Fusarium culmorum*. *Journal of Phytopathology* 148(5): 263-275.
- Kang, Z. and H. Buchenauer. 2002. Immunocytochemical localization of cell wall-bound thionins and hydroxyproline-rich glycoproteins in *Fusarium culmorum*-infected wheat spikes. *Journal of Phytopathology* 151(3): 120-129.
- Kazarian S.G. and K.L.A. Chan. 2006. Applications of atr-ftir spectroscopic imaging to biomedical samples. *Biochimica et Biophysica Acta*. 1758(7): 858-867.
- Kleeman, L. 1996. Understanding and applying kalman filtering. Sydney, Australia: Curtin University of Technology.
- Kos, G., H. Lohninger, and R. Krska. 2002. Fourier transform mid-infrared spectroscopy with attenuated total reflection (ft-ir/atr) as a tool for the detection of *Fusarium* fungi on maize. *Vibrational Spectroscopy* 29(1-2): 115-119.

- Kos, G., H. Lohninger, B. Mizaikoff, and K. Rudolf. 2007. Optimisation of a sample preparation procedure for screening of fungal infection and assessment of deoxynivalenol content in maize using mid-infrared attenuated total reflection spectroscopy. *Food Additives and Contaminants* 24(7): 721-729.
- Laidler, K.J., Meiser, J.H., and B.C. Sanctuary. 2003. *Physical Chemistry*, 4th edition. Boston, MA: Houghton Mifflin Company.
- Lattanzio, M.T.V., M. Pascale, and A. Visconti. 2009. Current analytical methods for trichothecene mycotoxins in cereals. *Trends in Analytical Chemistry* 28(6): 758-768.
- Lellmann, J., J. Balzer, A. Rieder, and J. Beyerer. 2008. Shape of specular reflection and optical flow. *International Journal of Computer Vision* 80(2): 226-241.
- Leung, A.K., F. Chau, and J. Gao. 1998. A review on the applications of wavelet transform techniques in chemical analysis: 1989-1997. *Chemometrics and Intelligent Laboratory Systems*. 43(1-2):165-184
- Lewis, E.N., P.J. Treado, R.C. Reeder, G.M. Story, A.E. Dowrey, C. Marcott, and I.W. Levin. 1995. Fourier transform spectroscopic imaging using an infrared focal-plane array detector. *Analytical Chemistry* 67(19):3377-3381.
- Liao, C., M.D. Piercey-Normore, J.L. Sorensen and K. Gough. 2010. In situ imaging of usnic acid in selected *Cladonia* spp. by vibrational spectroscopy. *Analyst* (135): 3242-3248.
- Life Science Weekly. 2012. Findings from University of Manchester Broaden Understanding of Algorithms.
<http://go.galegroup.com.proxy2.lib.umanitoba.ca/ps/i.do?id=GALE%7CA284848170&v=2.1&u=univmanitoba&it=r&p=HRCA&sw=w> (2012/18/04)

- Liu, S., S. Kokot, and G. Will. 2010. Photochemistry and chemometrics – an overview. *Journal of Photochemistry and Photobiology C: Photochemistry Reviews* 10(4): 159-172.
- Mahmoud, M.A. 2008. Phase 1 analysis of multiple linear regression profiles. *Communications in Statistics – Simulation and Computation* 37(10): 2016-2130.
- Maleki, M.R., A.M. Mouazen, H. Ramon, J.D. Baerdemaeker. 2006. Multiplicative scatter correction during on-line measurement with near infrared spectroscopy. *Biosystems Engineering* 96(3): 427-433.
- Marcott, C., R.C. Reeder, J.A. Sweat, D.D. Panzer, and D.L. Wetzel. 1999. FT-IR spectroscopic imaging microscopy of wheat kernels using a mercury-cadmium-telluride focal-plane array detector. *Vibrational Spectroscopy* 19(1): 123-129.
- Marill, K.A. 2003. Advanced statistics: linear regression part ii: multiple linear regression. *Academic Emergency Medicine* 11(1): 94-102.
- Martin, F.L., J.G. Kelly, V. Llabjani, P.L. Martin-Hirsch, I.I. Patel, J. Trevisan, N.J. Fullwood and M.J. Walsh. 2010. Distinguishing cell types or populations based \ on the computational analysis of their infrared spectra. *Nature Protocols* 5(11): 1748-1760.
- Martin-Gil, J., G. Palacios-Leble, P.M. Ramos and F.J. Martin-Gil. 2007. Analysis of celtiberian protective paste and its possible use by arevaci warriors. *Journal of Interdisciplinary Celtic Studies* 5: 63-76.
- McMullen, M., S. Zhong and S. Neate. 2008. Fusarium Head Blight (Scab) of Small Grains. <http://www.ag.ndsu.edu/pubs/plantsci/smgrains/pp804w.htm>. (20/04/2012)

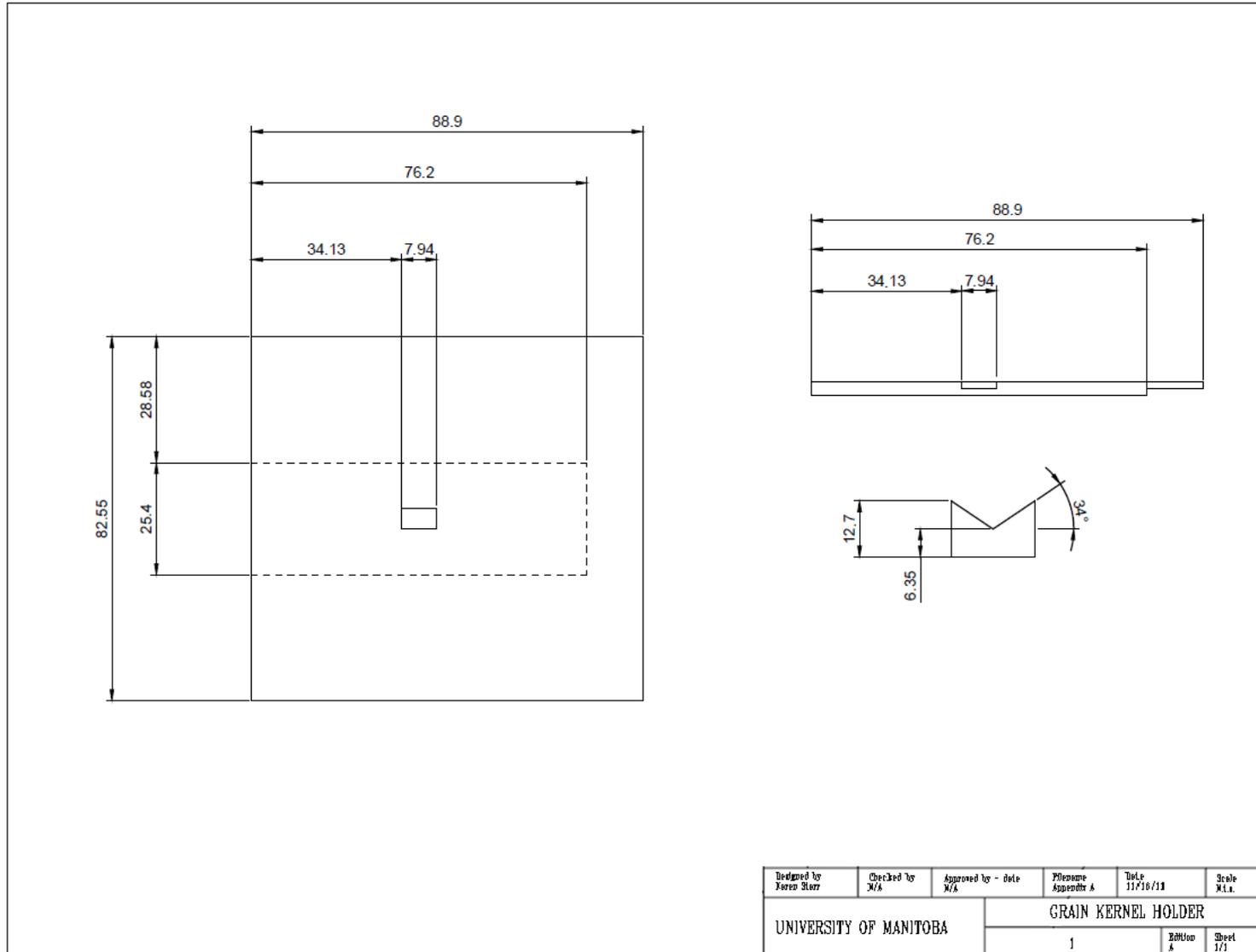
- Meyers, J.L., A.D. Well, and R.F. Lorch Jr. 2010. *Research Design and Statistical Analysis*, 3rd edition. New York, NY: Routledge.
- Milosevic, M. and S.L. Berets. 2002. A review of ft-ir diffuse reflectance sampling considerations. *Applied Spectroscopy Reviews* 37(4): 347-364.
- Murray, I. 2003. *Scattered Information: Philosophy and Practice of Near Infrared Spectroscopy*. Scottish Agricultural College, Aberdeen, UK.
- Naumann, A., M. Navarro-Gonzalez, S. Peddireddi, U. Kües and A. Polle. 2005. Fourier transform infrared microscopy and imaging: detection of fungi in wood. *Fungal Genetics and Biology* 42(10): 829-835.
- Nie, M., J. Luo, M. Xiao, J. Chen, K. Bao, W. Zhang, J. Chen, B. Li. 2007. Structural differences between fusarium strains investigated by ft-ir spectroscopy *Biochemistry* 72(1): 61-67.
- Osborne, B.G., T. Fearn and P.H. Hindle. 1993. *Practical NIR Spectroscopy with Applications in Food and Beverage Analysis*, 2nd edition. Essex, England: Longman Scientific & Technical; New York: John Wiley and Sons.
- Osborne, L.E. and J.M. Stein. 2007. Epidemiology of fusarium head blight on small-grain cereals. *International Journal of Food Microbiology* 119(1-2): 103-108.
- Pedrotti, L.S. 2001. Basic physical optics. In *Fundamentals of Photonics*, ed. B.E.A. Saleh and M.C. Teich, 152-154. Hoboken, NJ: John Wiley & Sons, Inc.
- Perkin Elmer. 2005. FT-IR spectroscopy attenuated total reflectance. 1-5. Perkin Elmer Life and Analytical Sciences. Shelton, CT, USA.
- Pike Technologies. 2000. Horizontal Attenuated Total Reflectance Accessory: Installation and User Guide. Madison, WI: Pike Technologies, Inc.

- Pinson-Gadais, L., C. Barreau, M. Chaurand, S. Gregoire, M. Monmarson, and F. Richard-Forget. 2007. Distribution of toxigenic fusarium spp. and mycotoxin production in milling fractions of durum wheat. *FEMS Microbiology Letters* 258(1): 102-107.
- Qin, S.J. 1993. Partial least squares regression for recursive system identification. *Decision and Control, 1993, Proceedings of the 32nd IEEE Conference* 3(15-17): 2617-2622.
- Reddy, V. U. 1998. On fast fourier transform: a popular tool for spectrum analysis. *Resonance* 3(10): 79-88.
- Salman, A., A. Pomerantz, L. Tsrer, I. Lapidot, A. Zwielly, R. Moreh, S. Mordechai and M. Huleihel. 2011. Distinction of fusarium oxysporum fungal isolates (strains) using ftir-atr spectroscopy and advanced statistical methods. *Analyst* 136: 988-995
- Salzer, R. and H.W. Siesler. 2009. *Infrared and Raman Spectroscopic Imaging*. Verlag, Germany: Wiley-VCH.
- Santos, C., M.E. Fraga, Z. Kozakiewicz, and N. Lima. 2010. Fourier transform infrared as a powerful technique for the identification and characterization of filamentous fungi and yeasts. *Research in Microbiology* 161(2): 168-175.
- Savitzky, A. and M.J.E. Golay. 1964. Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry* 36(8): 1627-1639.
- Shahin, M.A. and S.J. Symons. 2011. Detection of fusarium damaged kernels in canadian western red spring wheat using visible/near-infrared hyperspectral imaging and principal component analysis. *Computers and Electronics in Agriculture* 75(1): 107-112.

- Singh, P., H.C. Andola, M.S.M. Rawat, G.J.N. Pant, and V.K. Purohit. 2011. Fourier transform infrared (ft-ir) spectroscopy in an-overview. *Research Journal of Medicinal Plant* 5(2):127-135.
- Sinha, R.C. and M.E. Savard. 1997. Concentration of deoxynivalenol in single kernels and various tissues in wheat heads. *Canadian Journal of Plant Pathology* 19(1): 8-12.
- Statistics Canada. 2010. Field crop reporting series: november estimates of production of principal field crops. 22-002-X. Ottawa, Ontario: Statistics Canada, Agriculture Division.
- Szeghalmi, A., S. Kaminskyj and K.M. Gough. 2006. A synchrotron ftir microspectroscopy investigation of fungal hyphae grown under optimal and stressed conditions. *Analytical and Bioanalytical Chemistry* 387: 1779-1789.
- Tobias, R.D. 1999. An introduction to partial least squares regression. 1-8. Cary, NC: SAS Institute Inc.
- Varian, Inc. 2008. 660/670/680-IR: ft-ir spectroscopy solutions. S1-1196. Mississauga, ON: Agilent Technologies.
- Xu, X.M. and A.M. Berrie. 2005. Epidemiology of mycotoxigenic fungi associated with fusarium ear blight and apple blue mould: a review. *Food Additives and Contaminants* 22(4): 290-301
- Xu, X. and P. Nicholson. 2009. Community ecology of fungal pathogens causing wheat head blight. *Annual Review of Phytopathology* 47: 83-103

- Walczak, B., and D.L. Massart. 2000. Chapter 15 calibration in wavelet domain. *Data Handling in Science and Technology* 22:323-349.
- Walker, A.M., P. Yu, C.R. Christensen, D.A. Christensen, and J.J. McKinnon. 2009. Fourier transform infrared microspectroscopic analysis of the effects of cereal type and variety within a type of grain on structural makeup in relation to rumen degradation kinetics. *Journal of Agricultural and Food Chemistry* 57(15): 6871-6878.
- Wang, W., and J. Paliwal. 2007. Near-infrared spectroscopy and imaging in food quality and safety. *Sensing and Instrumentation for Food Quality and Safety* 1(4):193-207.
- Yeniay, Ö., and A. Götaş. 2002. A comparison of partial least squares regression with other prediction methods. *Hacettepe Journal of Mathematics and Statistics*. 31:99-111.

APPENDIX A: DESIGN OF A GRAIN KERNEL HOLDER



Designed by Keren Slorz	Checked by N/A	Approved by - date N/A	Reference Appendix A	Date 11/18/11	Scale N.L.
UNIVERSITY OF MANITOBA			GRAIN KERNEL HOLDER		
			1	Edition A	Sheet 1/1

APPENDIX B: CANADIAN GRAIN COMMISSION'S OFFICIAL GRAIN GRADING GUIDE

Grade name	Standard of quality					Foreign material					Total %
	Minimum test weight kg/hl (g/0.5 L)	Variety	Min. hard vitreous kernels %	Min. protein %	Degree of soundness	Ergot %	Excreta %	Matter other than cereal grains %	Sclerotinia %	Stones %	
No.1 CWS	75 (365)	Any variety of the class CWS designated as such by order of the Commission	65	10	Reasonably well matured, reasonably free from damaged kernels	0.01	0.01	0.2	0.01	0.03	0.6
No. 2 CWS	72 (350)	Any variety of the class CWS designated as such by order of the Commission	No minimum	No minimum	Fairly well matured, may be moderately bleached or frost-damaged, reasonably free from severely damaged kernels	0.02	0.01	0.3	0.02	0.03	1.2

Grade name	Standard of quality					Foreign material					
	Minimum test weight kg/hl (g/0.5 L)	Variety	Minimum hard vitreous kernels %	Minimum protein %	Degree of soundness	Ergot %	Excreta %	Matter other than cereal grains %	Sclerotinia %	Stones %	Total %
No. 3 CWR S	69 (335)	Any variety of the class CWR S designated as such by order of the Commission	No minimum	No minimum	May be frost-damaged, immature or weather-damaged, moderately free from severely damaged kernels	0.04	0.015	0.5	0.04	0.06	2.4
No. 4 CWR S	68 (330)	Any variety of the class CWR S designated as such by order of the Commission	No minimum	No minimum	May be severely frost-damaged, immature or weather-damaged, moderately free from other severely damaged kernels	0.04	0.015	0.5	0.04	0.06	2.4

Grade name	Standard of quality					Foreign material					
	Minimum test weight kg/hl (g/0.5 L)	Variety	Minimum hard vitreous kernels %	Minimum protein %	Degree of soundness	Ergot %	Excreta %	Matter other than cereal grains %	Sclerotinia %	Stones %	Total %
CW Feed	65 (315)	Any class or variety of wheat excluding amber durum and General purpose	No minimum	No minimum	Reasonably sweet, excluded from other grades of wheat on account of damaged kernels	0.1	0.03	1	0.1	0.1	10
Grade, if specs for CW Feed not met	Wheat, Sample CW Account Light Weight	N/A	N/A	N/A	N/A	Wheat, Sample CW Account Ergot	Wheat, Sample CW Account Excreta	Wheat, Sample CW Account Admixture	Wheat, Sample CW Account Admixture	2.5% or less – Wheat, Rejected grade, Account Stones Over 2.5% – Wheat, Sample	See Mixed grain

Grade name	Wheats of other classes or varieties (*)		Artificial stain, no residue %	Dark, Immature %	Degermed %	Fireburn t %	Fusarium damage %	Grass green %	Grasshopper, army worm %	Heated	
	Contrasting classes %	Total %								Binburnt severely mildewed rotted, mouldy %	Total %
No.1 CWRS	<u>0.75</u>	<u>2.3</u>	Nil	1	4	Nil	<u>0.25</u>	<u>0.75</u>	1	0.005	0.05
No. 2 CWRS	<u>2.3</u>	<u>4.5</u>	0.05	<u>2.5</u>	7	Nil	0.8	2	3	0.02	0.4
No. 3 CWRS	<u>3.8</u>	<u>7.5</u>	0.1	10	13	Nil	1.5	10	8	0.03	1
No. 4 CWRS	<u>3.8</u>	<u>7.5</u>	0.1	10	13	Nil	1.5	10	8	0.03	1
CW Feed	No limit-but not more than 10% amber durum and/or General		2	No limit	No limit	2	4	No limit	No limit	<u>2.5</u>	<u>2.5</u>

purpose

Grade name	Wheats of other classes or varieties (*) Contrasting classes %/Total %	Artificial stain, no residue %	Dark, Immature %	Degermed %	Fireburn t %	Fusarium damage %	Grass green %	Grasshopper, army worm %	Heated Binburnt severely mildewed rotted, mouldy %/ Total%
Grade, if specs for CW Feed not met	Over 10% amber durum and/or General purpose – Wheat, Sample CW Account Admixture	Wheat, Sample CW Account Stained Kernels	N/A	N/A	Wheat, Sample CW Account Fireburn t	Wheat, Sample CW Account Fusarium Damage Over 10%- Wheat, Commercial Salvage	N/A	N/A	Wheat, Sample CW Account Heated

Grade name	Natural stain %	Pink %	Sawfly, midge %	Shrunken and broken (**)			Smudge and blackpoint		Sprouted	
				Shrunken %	Broken %	Total %	Smudge %	Total %	Severely sprouted %	Total %
No.1 CWRS	0.5	<u>1.5</u>	2	4	5	7	0.3	10	0.1	0.5
No. 2 CWRS	2	5	5	4	6	8	1	20	0.2	1
No. 3 CWRS	5	10	10	4	7	9	5	35	0.3	3
No. 4 CWRS	5	10	10	4	7	9	5	35	0.5	5
CW Feed	No limit	No limit	No limit	No limit	13	No limit within broken tolerances	No limit	No limit	No limit	No limit
Grade, if specs for CW Feed not met	N/A	N/A	N/A	N/A	Sample Broken Grain	N/A	N/A	N/A	N/A	N/A

**APPENDIX C: SOIL ZONE, LOCATION, AND FUSARIUM DAMAGE
PERCENTAGE OF SAMPLES**

Soil Zone	Sample Number	Location	Grade	Fusarium Damage (%)
Black	1	Balcarres	CW Feed	1.10
Black	2	Balcarres	CW Feed	0.81
Black	3	Dauphin	No. 2 CWRS	0.80
Black	4	Elm Creek	CW Feed	3.00
Black	5	Elm Creek	No. 2 CWRS	1.00
Black	6	Elm Creek	No. 2 CWRS	1.35
Black	7	Elm Creek	No. 2 CWRS	1.17
Black	8	Elm Creek	No. 2 CWRS	0.90
Black	9	Elm Creek	No. 1 CWRS	0.70
Black	10	Elva	No. 3 CWRS	0.66
Black	11	Elva	No. 3 CWRS	0.80
Black	12	Morris	Sple Acct Fusarium	4.96
Black	13	Morris	CW Feed	1.97
Black	14	Morris	Sple Acct Fusarium	4.24
Black	15	Morris	Sple Acct Fusarium	6.12
Black	16	Morris	No. 1 CWRS	0.22
Black	17	Morris	No. 1 CWRS	0.17
Black	18	Morris	No. 1 CWRS	0.12
Black	19	Nicklen Siding	No. 3 CWRS	0.82
Black	20	Nicklen Siding	CW Feed	2.77
Black	21	Nicklen Siding	CW Feed	3.56
Black	22	Nicklen Siding	No. 3 CWRS	0.98
Black	23	Oakner	CW Feed	1.71
Black	24	Oakner	CW Feed	4.00
Black	25	Oakner	Sple Acct Fusarium	6.00
Black	26	Oakner	CW Feed	3.00
Black	27	Oakner	CW Feed	0.40
Black	28	Swan River	Sple Acct Fusarium	8.84
Black	29	Viking West	No. 4 CWRS	1.10
Black	30	Yorkton	No. 3 CWRS	1.20
Black	31	Yorkton	No. 3 CWRS	0.80
Black	32	Elva	No. 3 CWRS	0.70
Black	33	Nicklen Siding	No. 2 CWRS	0.50
Black	34	Nicklen Siding	No. 2 CWRS	0.71

Soil Zone	Sample Number	Location	Grade	Fusarium Damage (%)
Black	35	Swan River	Wheat Commerical Salvage	12.40
Black	36	Elva	No. 3 CWRS	0.49
Black	37	Elva	No. 2 CWRS	0.80
Black	38	Elva	No. 2 CWRS	0.04
Black	39	Elva	CW Feed	1.42
Black	40	Balcarres	No. 4 CWRS	0.70
Black	41	Balcarres	No. 3 CWRS	1.45
Black	42	Balcarres	No. 3 CWRS	0.72
Black	43	Balcarres	No. 3 CWRS	1.35
Black	44	Balcarres	No. 3 CWRS	1.04
Black	45	Elm Creek	No. 2 CWRS	0.71
Black	46	Morris	No. 3 CWRS	1.10
Black	47	Morris	No. 3 CWRS	1.10
Black	48	Nicklen Siding	No. 2 CWRS	0.31
Black	49	Nicklen Siding	CW Feed	2.60
Black	50	Balcarres	No. 3 CWRS	0.47
Black	51	Balcarres	No. 3 CWRS	1.13
Black	52	Balcarres	No. 3 CWRS	1.16
Black	53	Nicklen Siding	No. 3 CWRS	0.92
Black	54	Nesbitt	No. 2 CWRS	0.45
Black	55	Oakner	No. 2 CWRS	0.63
Black	56	Oakner	No. 2 CWRS	0.59
Black	57	Elm Creek	No. 2 CWRS	0.90
Black	58	Elm Creek	No. 2 CWRS	0.57
Black	59	Elva	No. 3 CWRS	1.04
Black	60	Nesbitt	No. 2 CWRS	0.24
Black	61	Nesbitt	No. 2 CWRS	0.42
Black	62	Nicklen Siding	No. 2 CWRS	0.73
Black	63	Oakner	No. 2 CWRS	0.35
Black	64	Oakner	No. 3 CWRS	0.99
Black	65	Oakner	No. 2 CWRS	0.89
Black	66	Oakner	No. 3 CWRS	0.55
Black	67	Oakner	No. 3 CWRS	0.66
Black	68	Oakner	No. 3 CWRS	0.98
Black	69	Yorkton	Sple Acct Fusarium	4.67
Black	70	Oakner	No. 3 CWRS	1.28
Black	71	Oakner	No. 3 CWRS	0.62
Black	72	Oakner	No. 2 CWRS	0.69
Black	73	Oakner	No. 2 CWRS	0.27
Brown	1	Carseland	CW Feed	0.43

Soil Zone	Sample Number	Location	Grade	Fusarium Damage (%)
Brown	2	Kindersley	No. 4 CWRS	0.83
Brown	3	Lethbridge	No. 4 CWRS	0.31
Brown	4	Lethbridge	No. 2 CWRS	0.42
Brown	5	Lethbridge	CW Feed	0.30
Brown	6	Rosetown	Sple Acct Fusarium	6.94
Brown	7	Lethbridge	No. 2 CWRS	0.19
Brown	8	Rosetown	No. 4 CWRS	0.99
Brown	9	Lethbridge	No. 2 CWRS	0.33
Brown	10	Lethbridge	No. 3 CWRS	0.82
Brown	11	Lethbridge	No. 1 CWRS	0.25
Brown	12	Rosetown	CW Feed	0.70
Brown	13	Rosetown	CW Feed	2.93
Brown	14	Rosetown	CW Feed	3.31
Brown	15	Carseland	No. 3 CWRS	0.62
Dark Brown	1	Clavet	CW Feed	3.89
Dark Brown	2	Equity	CW Feed	0.87
Dark Brown	3	Equity	No. 3 CWRS	1.04
Dark Brown	4	Equity	CW Feed	0.42
Dark Brown	5	Equity	No. 3 CWRS	0.60
Dark Brown	6	Moosejaw	No. 3 CWRS	0.35
Dark Brown	7	Moosejaw	No. 3 CWRS	2.00
Dark Brown	8	Raymore	CW Feed	0.76
Dark Brown	9	Equity	CW Feed	3.33
Dark Brown	10	Moosejaw	No. 3 CWRS	0.72
Dark Brown	11	Equity	CW Feed	1.01
Dark Brown	12	Equity	CW Feed	1.00
Dark Brown	13	Moosejaw	No. 3 CWRS	0.43

**APPENDIX D: SAMPLES USED IN THE EXPERIMENT AS
DETERMINED BY A RANDOM NUMBER GENERATOR**

Soil Zone	Random Number Generated*	Location	Grade	Fusarium Damage (%)
Black	1	Balcarres	CW Feed	1.10
Black	4	Elm Creek	CW Feed	3.00
Black	5	Elm Creek	No. 2 CWRS	1.00
Black	6	Elm Creek	No. 2 CWRS	1.35
Black	8	Elm Creek	No. 2 CWRS	0.90
Black	13	Morris	CW Feed	1.97
Black	16	Morris	No. 1 CWRS	0.22
Black	23	Oakner	CW Feed	1.71
Black	23	Oakner	CW Feed	1.71
Black	27	Oakner	CW Feed	0.40
Black	27	Oakner	CW Feed	0.40
Black	27	Oakner	CW Feed	0.40
Black	29	Viking West	No. 4 CWRS	1.10
Black	39	Elva	CW Feed	1.42
Black	41	Balcarres	No. 3 CWRS	1.45
Black	41	Balcarres	No. 3 CWRS	1.45
Black	42	Balcarres	No. 3 CWRS	0.72
Black	58	Elm Creek	No. 2 CWRS	0.57
Black	63	Oakner	No. 2 CWRS	0.35
Black	63	Oakner	No. 2 CWRS	0.35
Black	68	Oakner	No. 3 CWRS	0.98
Black	68	Oakner	No. 3 CWRS	0.98
Black	69	Yorkton	Sple Acet Fusarium	4.67
Black	71	Oakner	No. 3 CWRS	0.62
Black	73	Oakner	No. 2 CWRS	0.27
Brown	1	Carseland	CW Feed	0.43
Brown	1	Carseland	CW Feed	0.43
Brown	1	Carseland	CW Feed	0.43
Brown	1	Carseland	CW Feed	0.43
Brown	1	Carseland	CW Feed	0.43
Brown	2	Kindersley	No. 4 CWRS	0.83
Brown	2	Kindersley	No. 4 CWRS	0.83
Brown	3	Lethbridge	No. 4 CWRS	0.31
Brown	3	Lethbridge	No. 4 CWRS	0.31
Brown	3	Lethbridge	No. 4 CWRS	0.31
Brown	3	Lethbridge	No. 4 CWRS	0.31

Soil Zone	Random Number Generated*	Location	Grade	Fusarium Damage (%)
Brown	4	Lethbridge	No. 2 CWRS	0.42
Brown	4	Lethbridge	No. 2 CWRS	0.42
Brown	5	Lethbridge	CW Feed	0.30
Brown	5	Lethbridge	CW Feed	0.30
Brown	5	Lethbridge	CW Feed	0.30
Brown	6	Rosetown	Sple Acct Fusarium	6.94
Brown	6	Rosetown	Sple Acct Fusarium	6.94
Brown	8	Rosetown	No. 4 CWRS	0.99
Brown	8	Rosetown	No. 4 CWRS	0.99
Brown	9	Lethbridge	No. 2 CWRS	0.33
Brown	10	Lethbridge	No. 3 CWRS	0.82
Brown	12	Rosetown	CW Feed	0.70
Brown	13	Rosetown	CW Feed	2.93
Brown	13	Rosetown	CW Feed	2.93
Dark Brown	1	Clavet	CW Feed	3.89
Dark Brown	2	Equity	CW Feed	0.87
Dark Brown	2	Equity	CW Feed	0.87
Dark Brown	3	Equity	No. 3 CWRS	1.04
Dark Brown	3	Equity	No. 3 CWRS	1.04
Dark Brown	4	Equity	CW Feed	0.42
Dark Brown	5	Equity	No. 3 CWRS	0.60
Dark Brown	5	Equity	No. 3 CWRS	0.60
Dark Brown	7	Moosejaw	No. 3 CWRS	2.00
Dark Brown	7	Moosejaw	No. 3 CWRS	2.00
Dark Brown	8	Raymore	CW Feed	0.76
Dark Brown	8	Raymore	CW Feed	0.76
Dark Brown	9	Equity	CW Feed	3.33
Dark Brown	9	Equity	CW Feed	3.33
Dark Brown	9	Equity	CW Feed	3.33
Dark Brown	9	Equity	CW Feed	3.33
Dark Brown	11	Equity	CW Feed	1.01
Dark Brown	11	Equity	CW Feed	1.01
Dark Brown	11	Equity	CW Feed	1.01
Dark Brown	11	Equity	CW Feed	1.01
Dark Brown	12	Equity	CW Feed	1.00
Dark Brown	12	Equity	CW Feed	1.00
Dark Brown	12	Equity	CW Feed	1.00
Dark Brown	13	Moosejaw	No. 3 CWRS	0.43
Dark Brown	13	Moosejaw	No. 3 CWRS	0.43

*Note: Random number generated was later sorted in ascending order

APPENDIX E: ATR PRESSURE POINT AND SPECTRAL IMAGES WITHIN THE ATR PRESSURE POINT

Figure E.0 Frequency slice depicting ATR pressure point

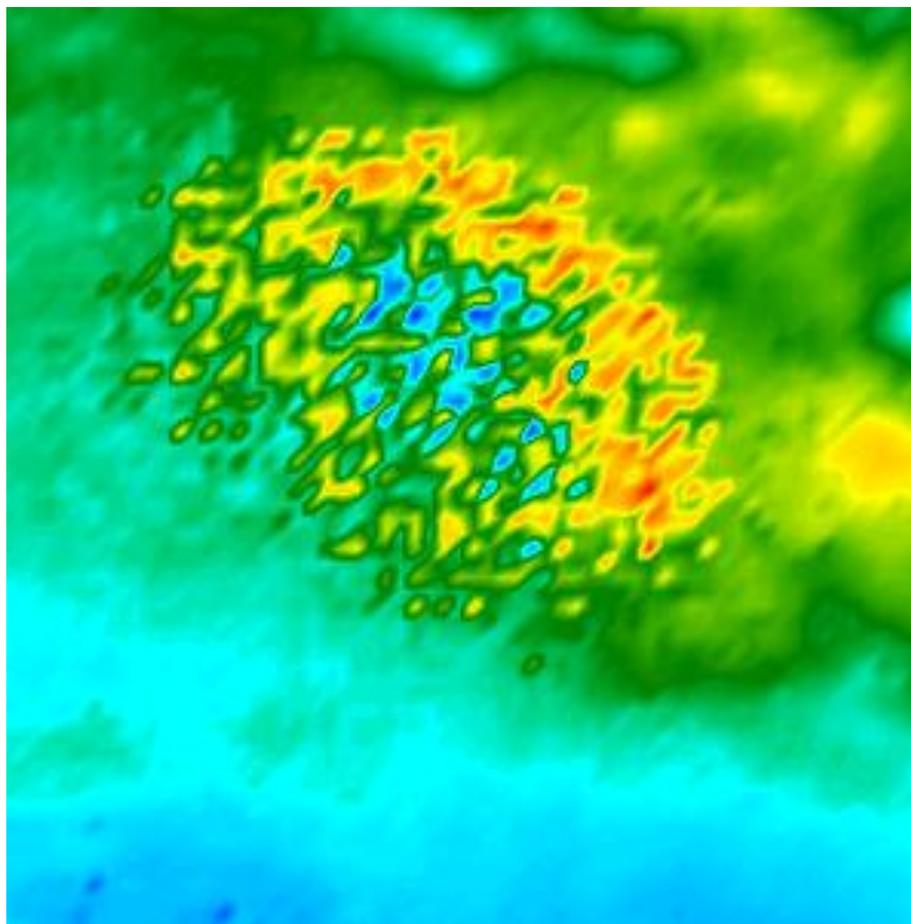
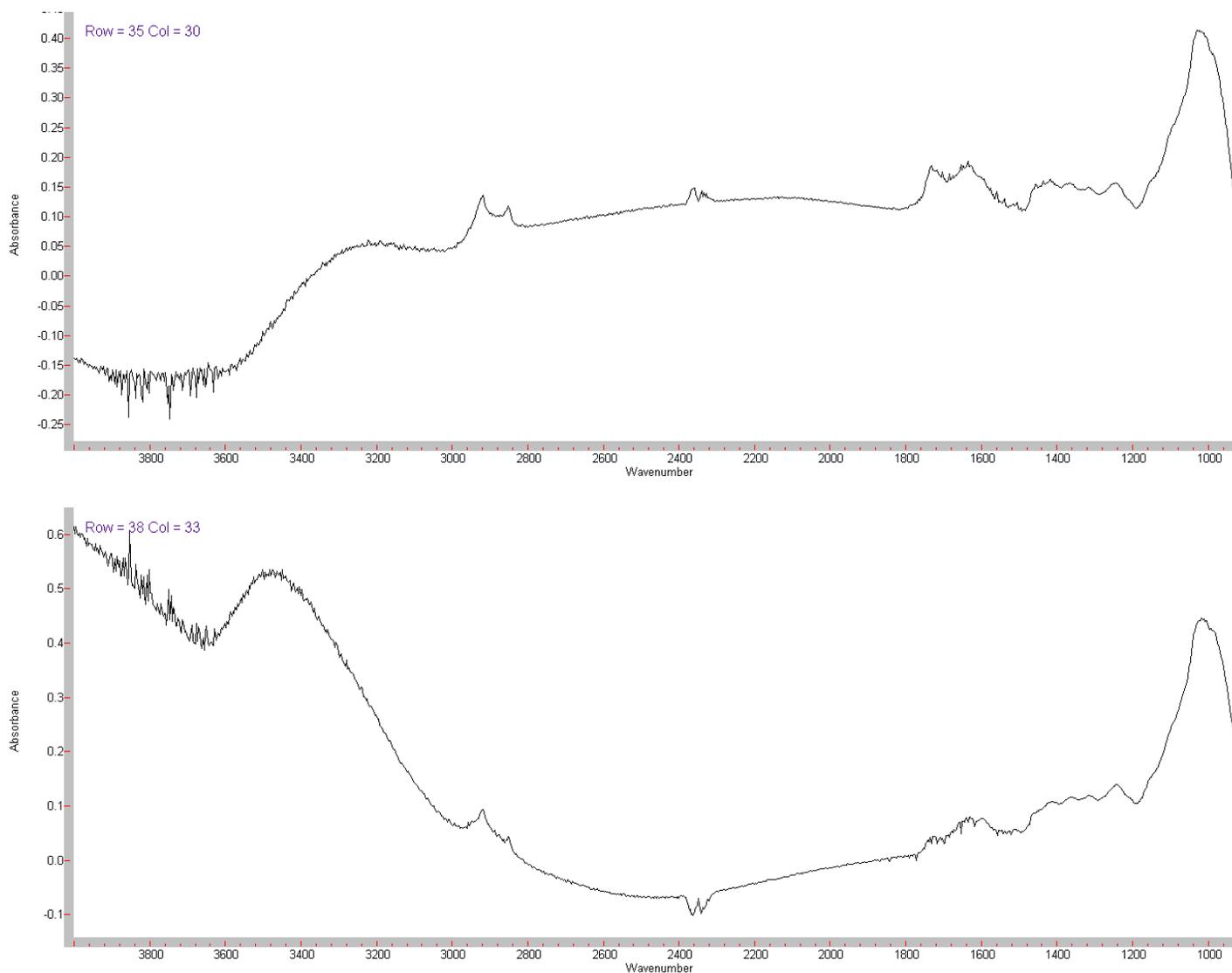


Figure E.1 Typical spectral images within the ATR pressure point exhibiting abnormal baselines



APPENDIX F: TYPICAL SPECTRAL IMAGES AND WHEAT KERNEL SURFACES

Figure F.0 Typical wheat kernel surface (a), typical spectrum (b), and details of functional groups (c) for an uninfected wheat kernel in the black soil zone

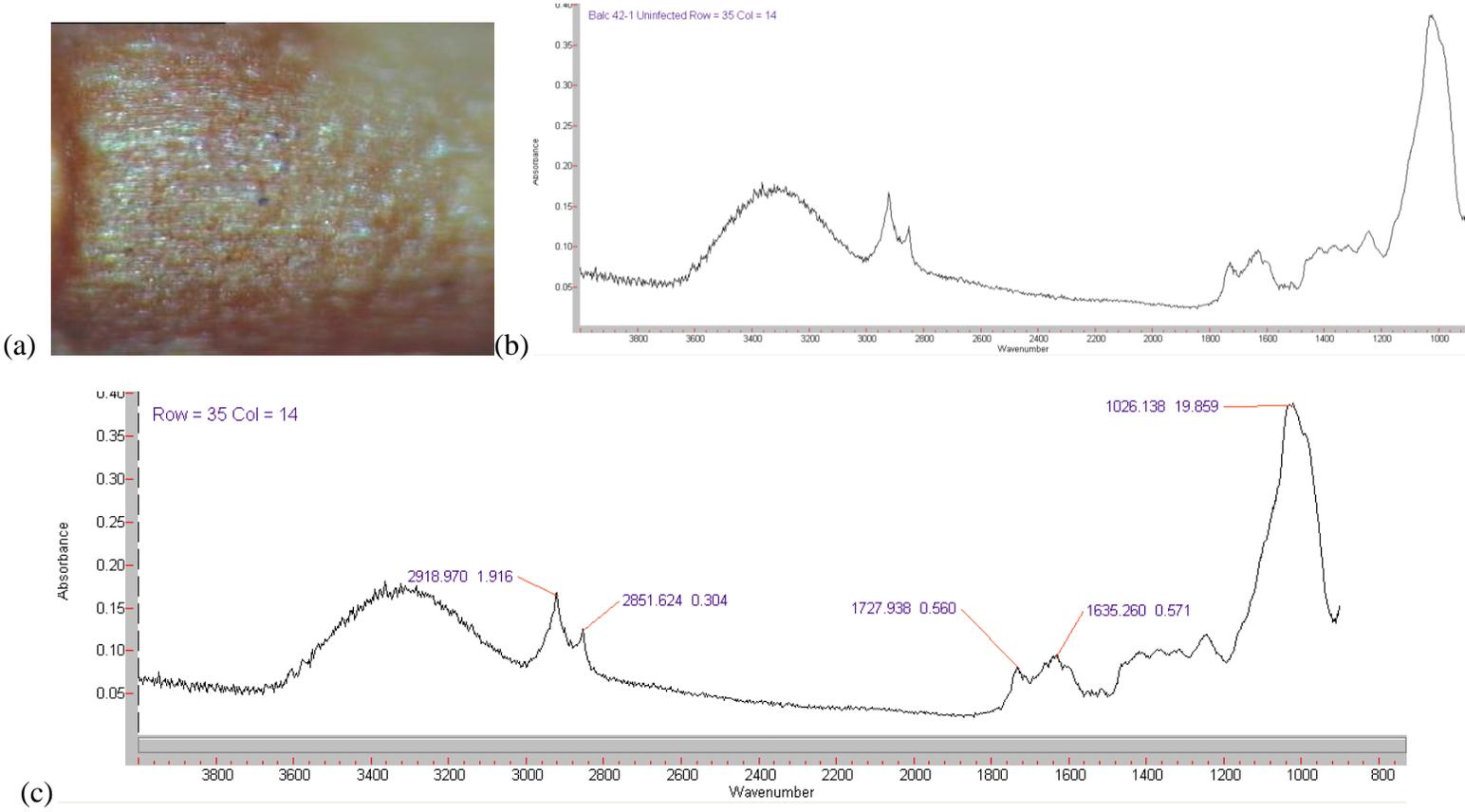
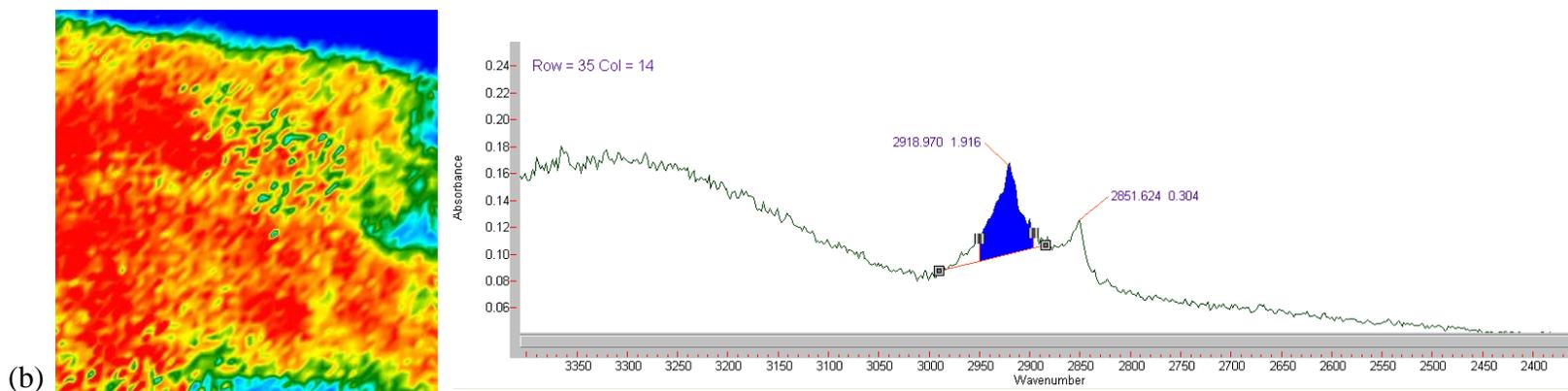
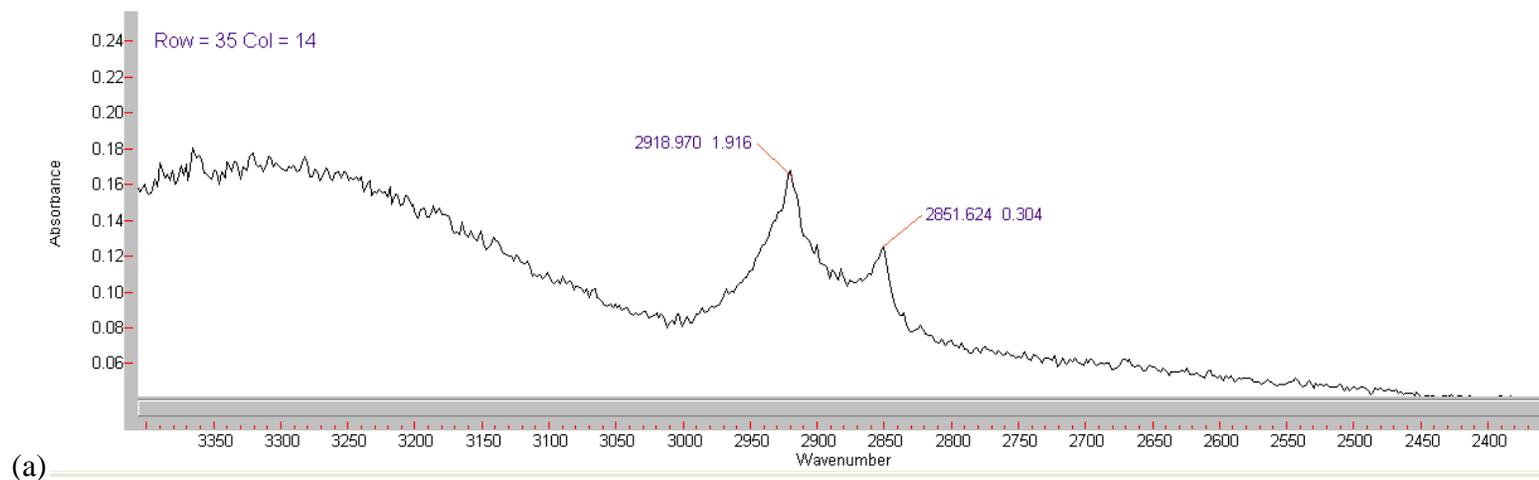
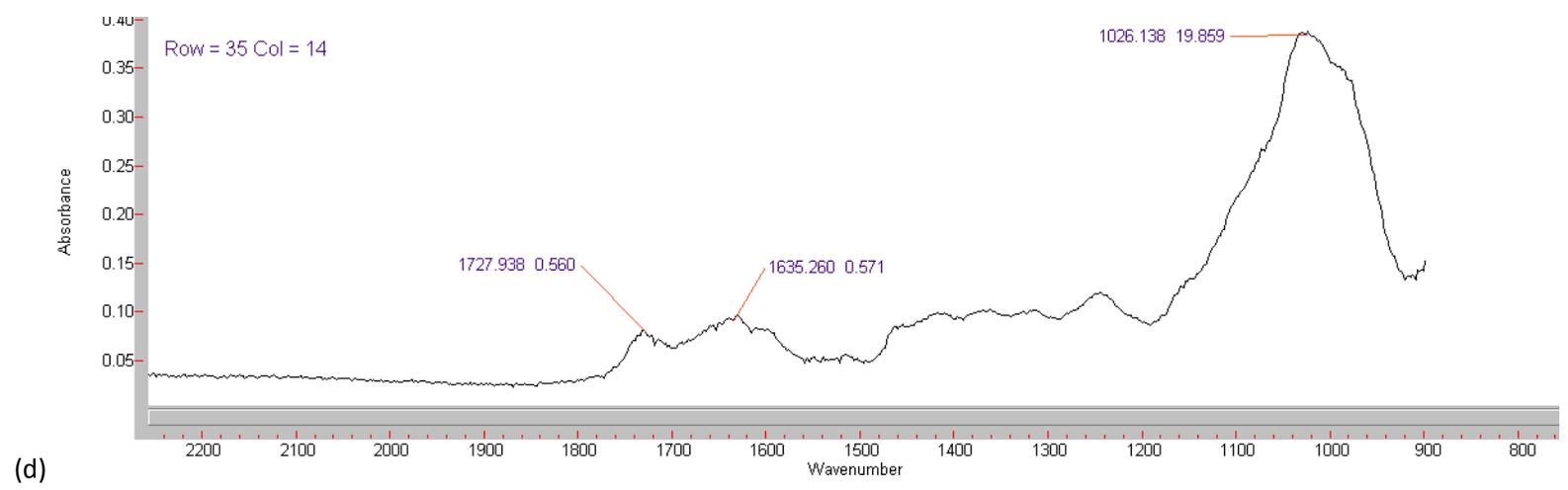
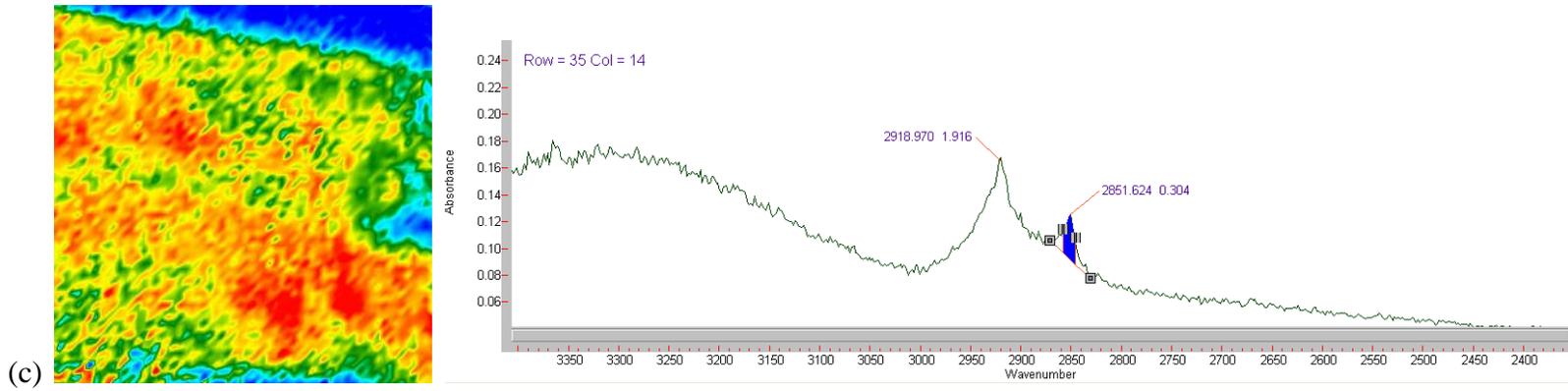


Figure F.1 Typical spectral functional groups and corresponding frequency slices for the CH₂ asymmetric (b) and CH₂ asymmetric (a) functional groups (a), the lipid carbonyl (e), protein amide one (f), and cellulose (g) functional groups (d) for an uninfected wheat kernel in the black soil zone





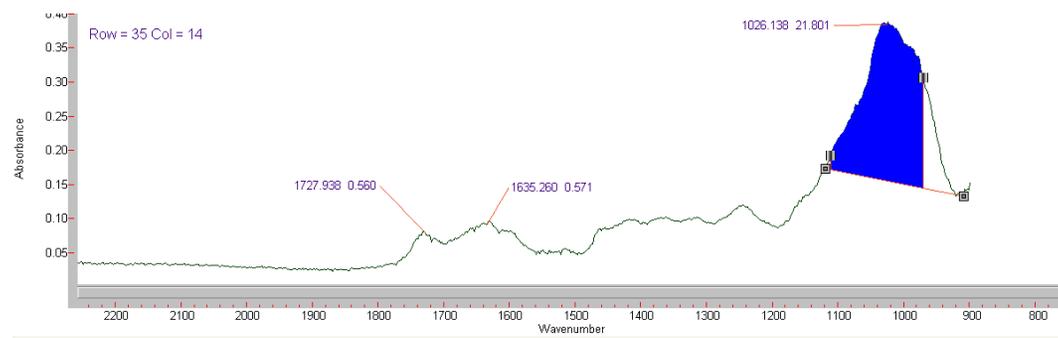
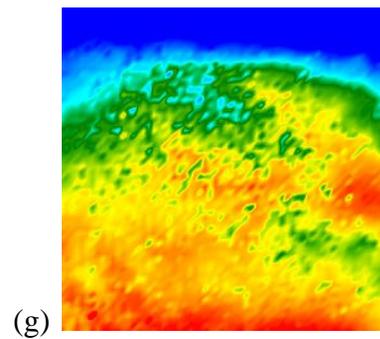
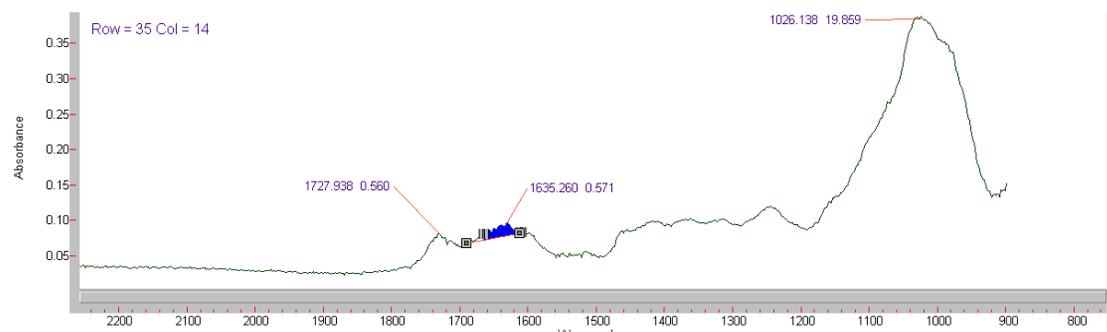
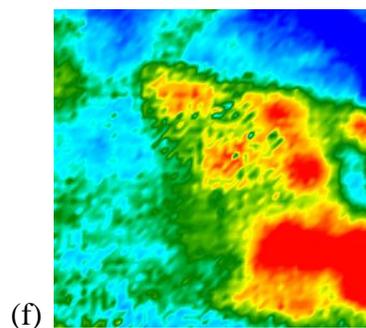
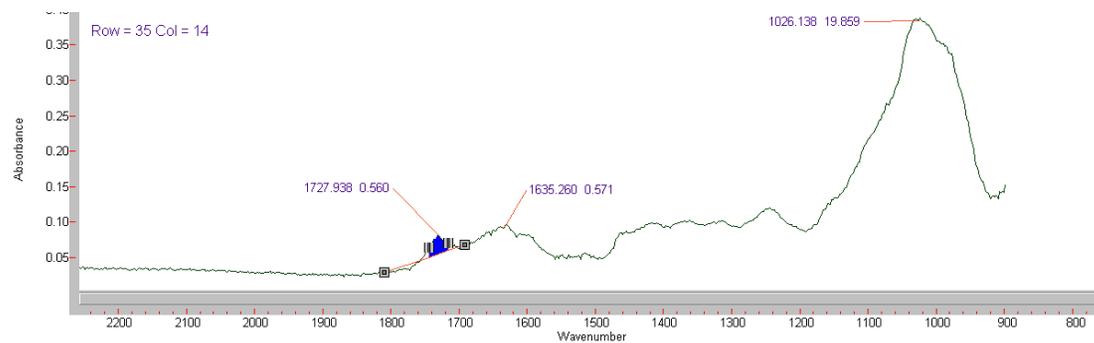
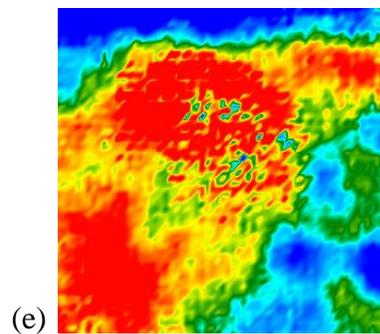


Figure F.2 Typical wheat kernel surface (a), typical spectrum (b), and details of functional groups (c) for an infected wheat kernel in the black soil zone

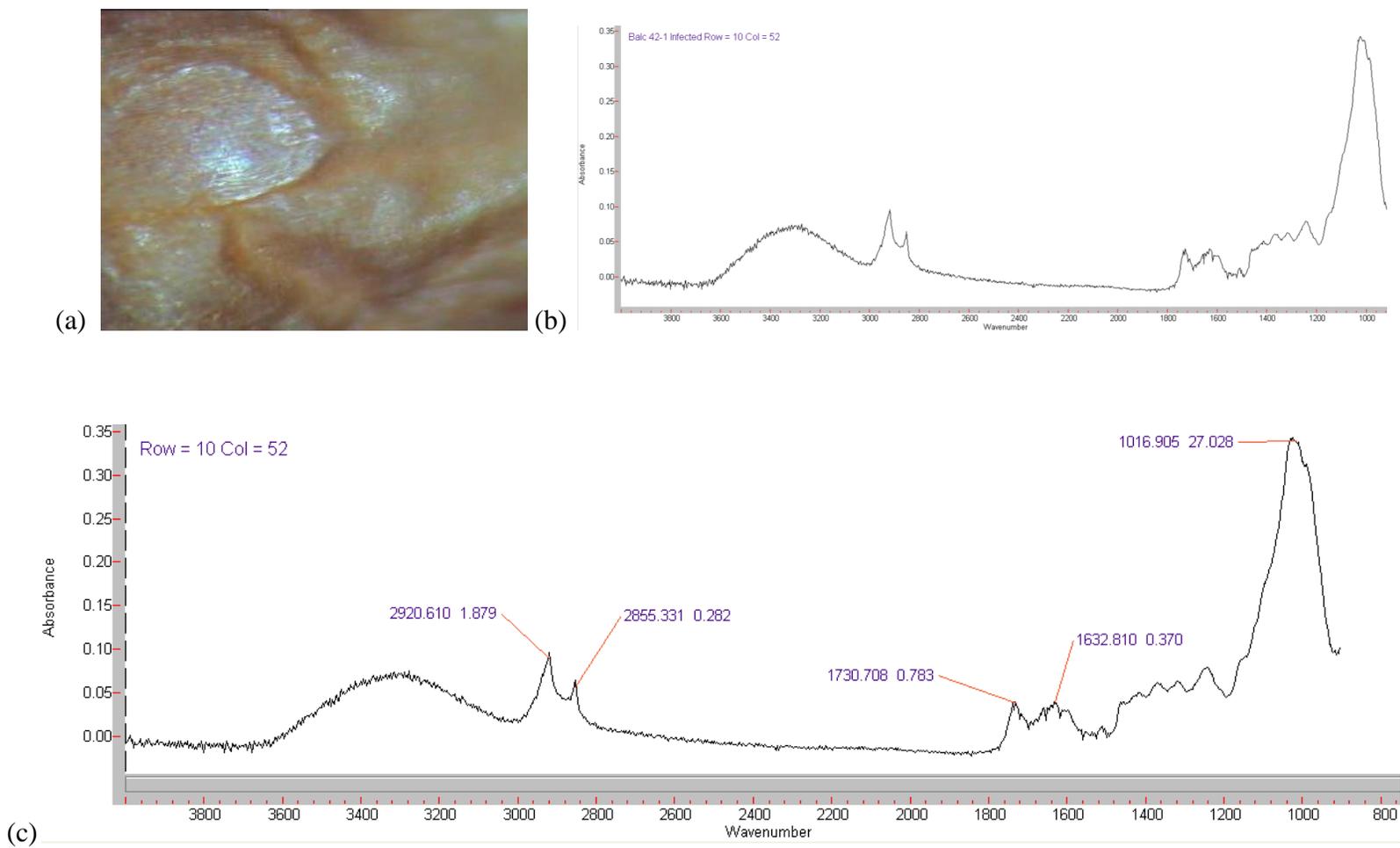
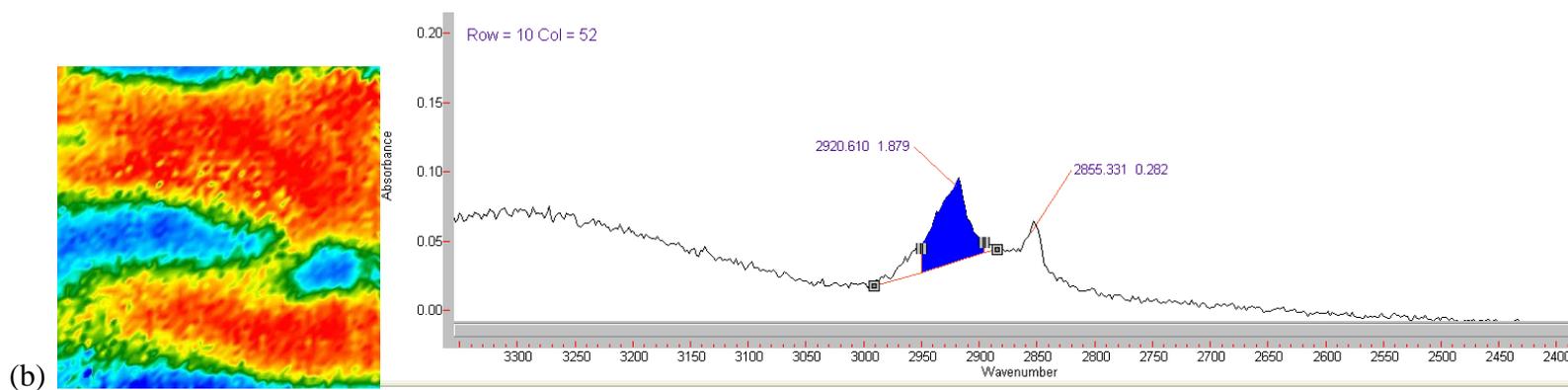
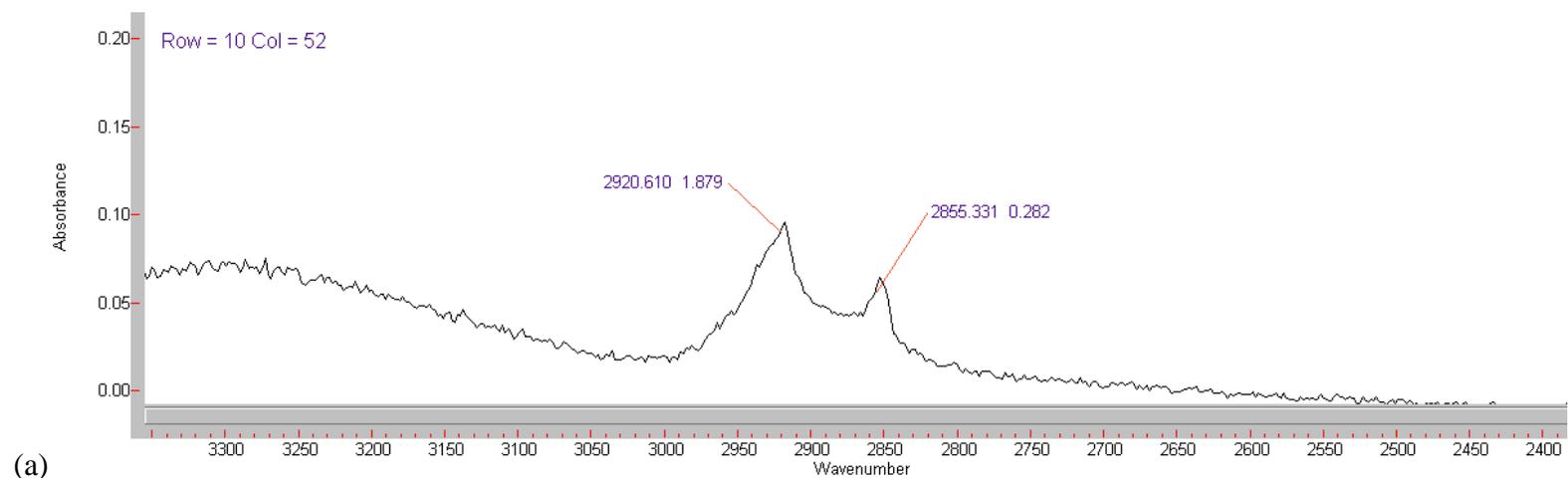
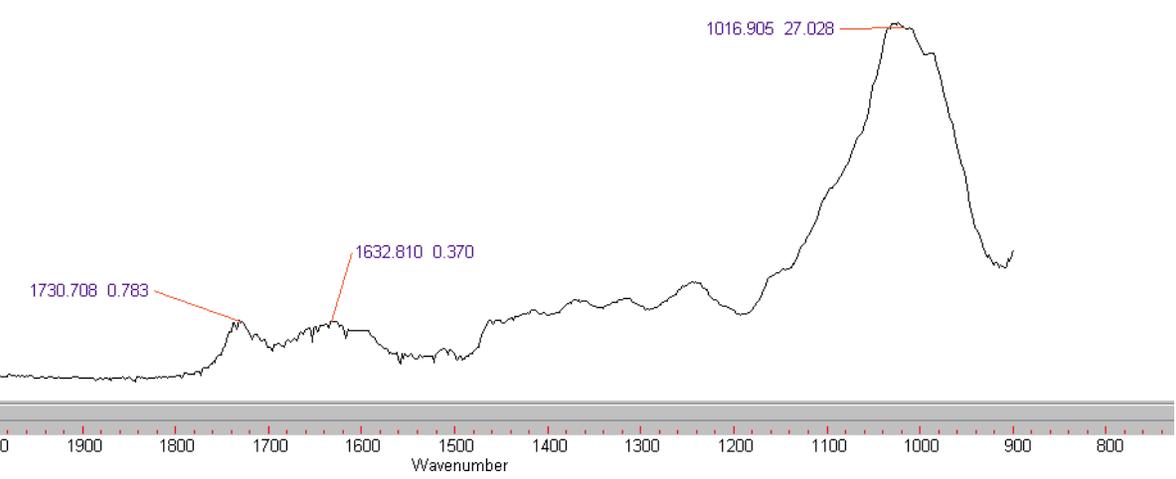
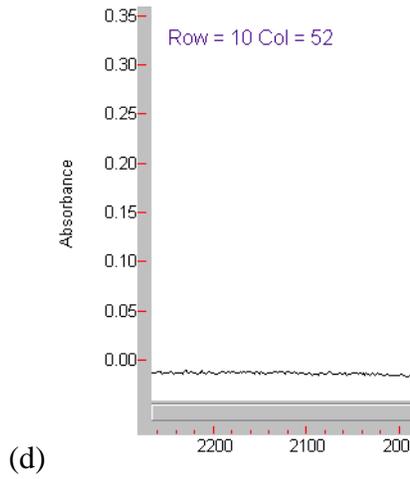
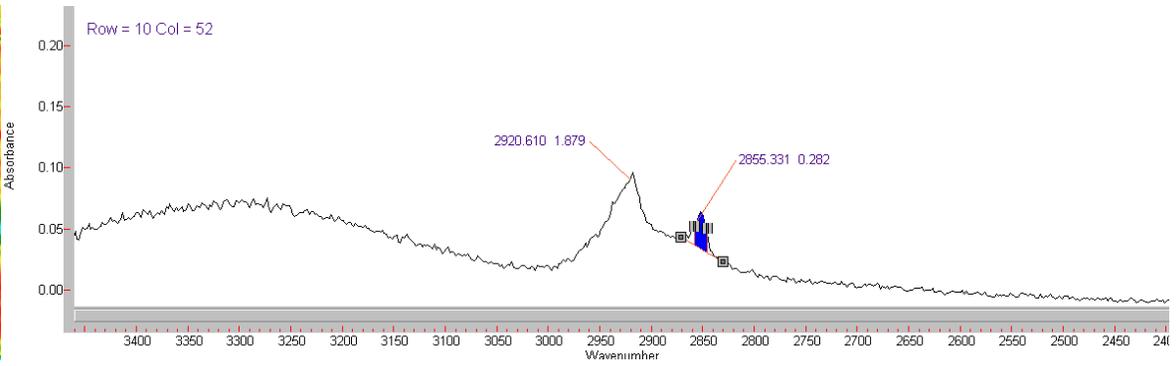
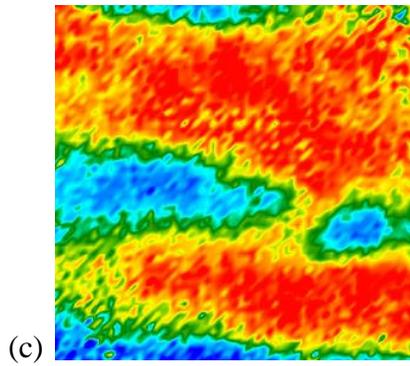


Figure F.3 Typical spectral functional groups and corresponding frequency slices for the CH₂ asymmetric (b) and CH₂ asymmetric (a) functional groups (a), the lipid carbonyl (e), protein amide one (f), and cellulose (g) functional groups (d) for an infected wheat kernel in the black soil zone





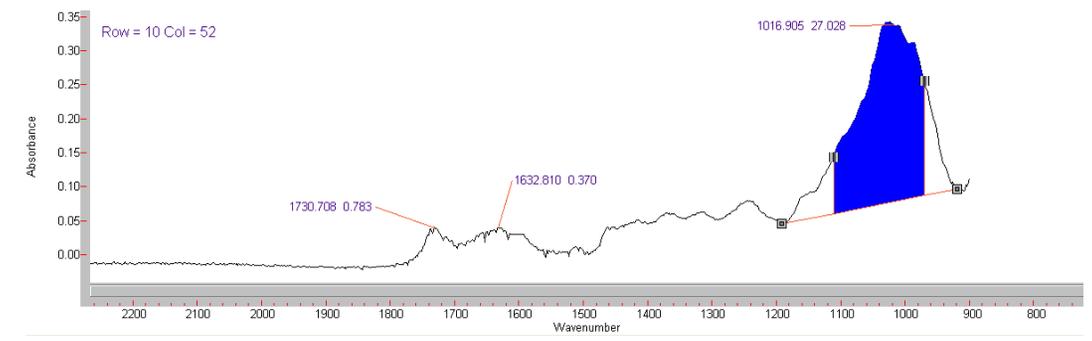
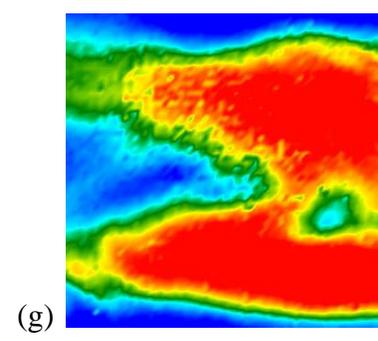
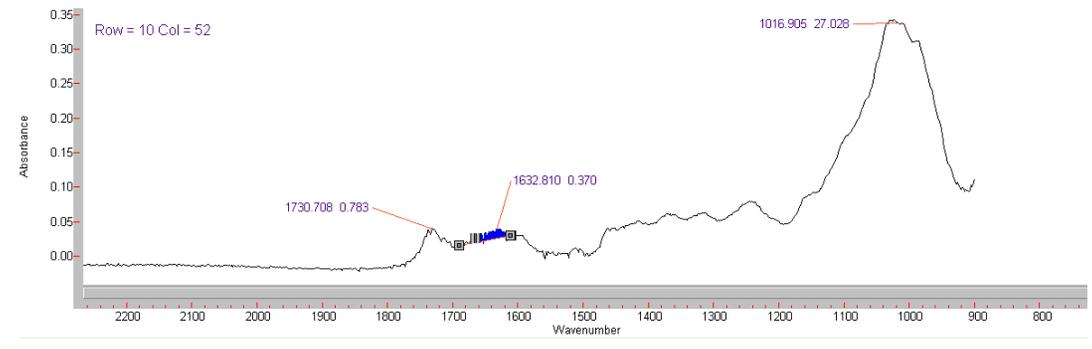
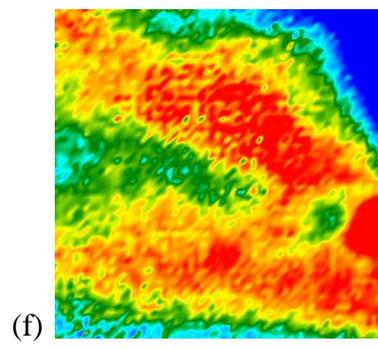
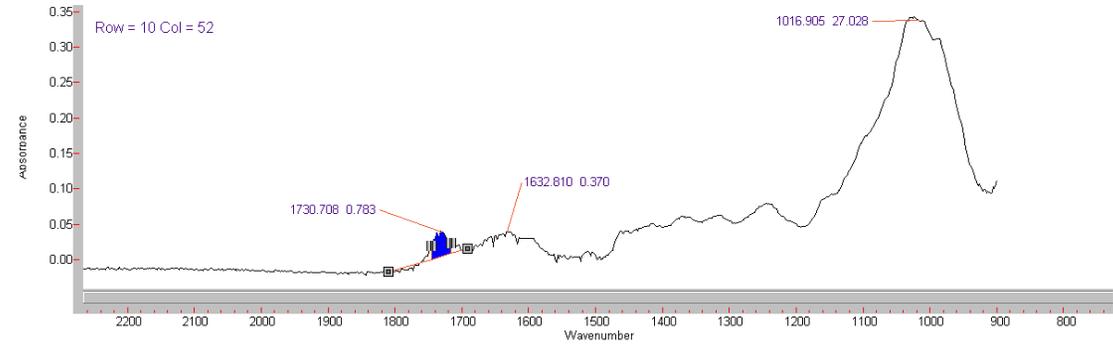
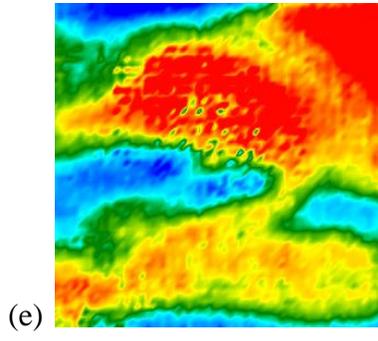


Figure F.4 Comparison of uninfected and infected spectra for kernels in the black soil zone

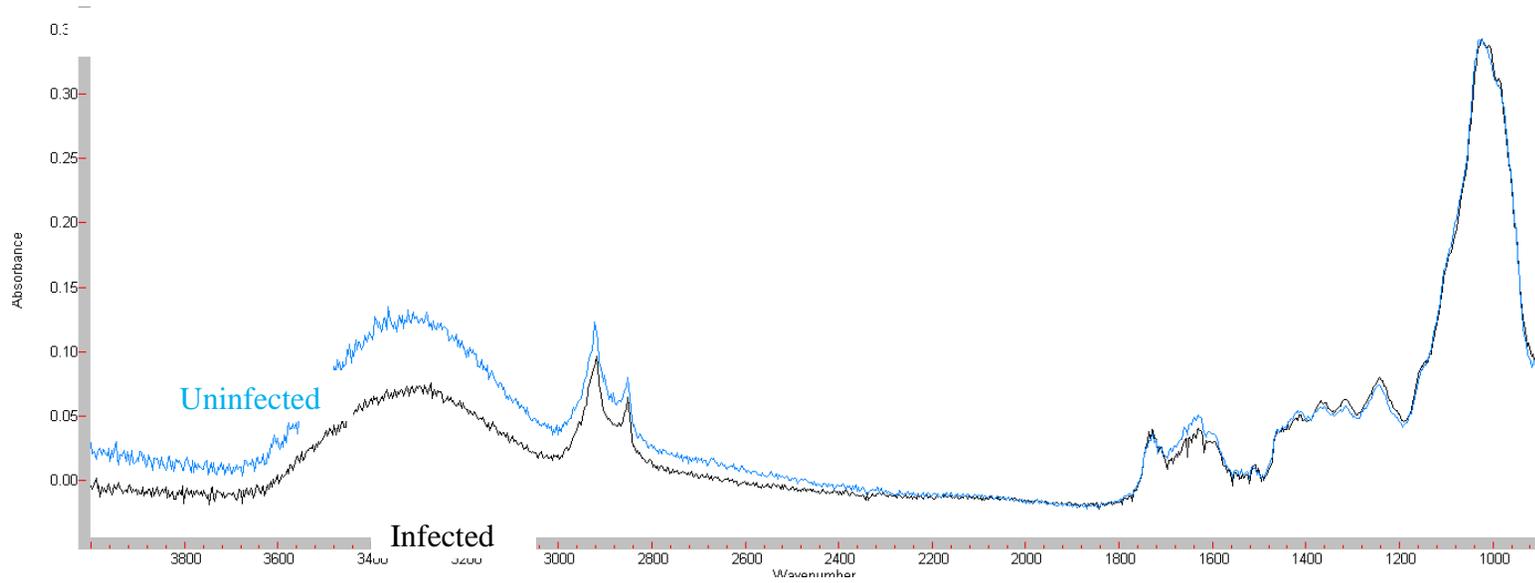


Figure F.5 Typical wheat kernel surface (a), typical spectrum (b), and details of functional groups (c) for an uninfected wheat kernel in the brown soil zone

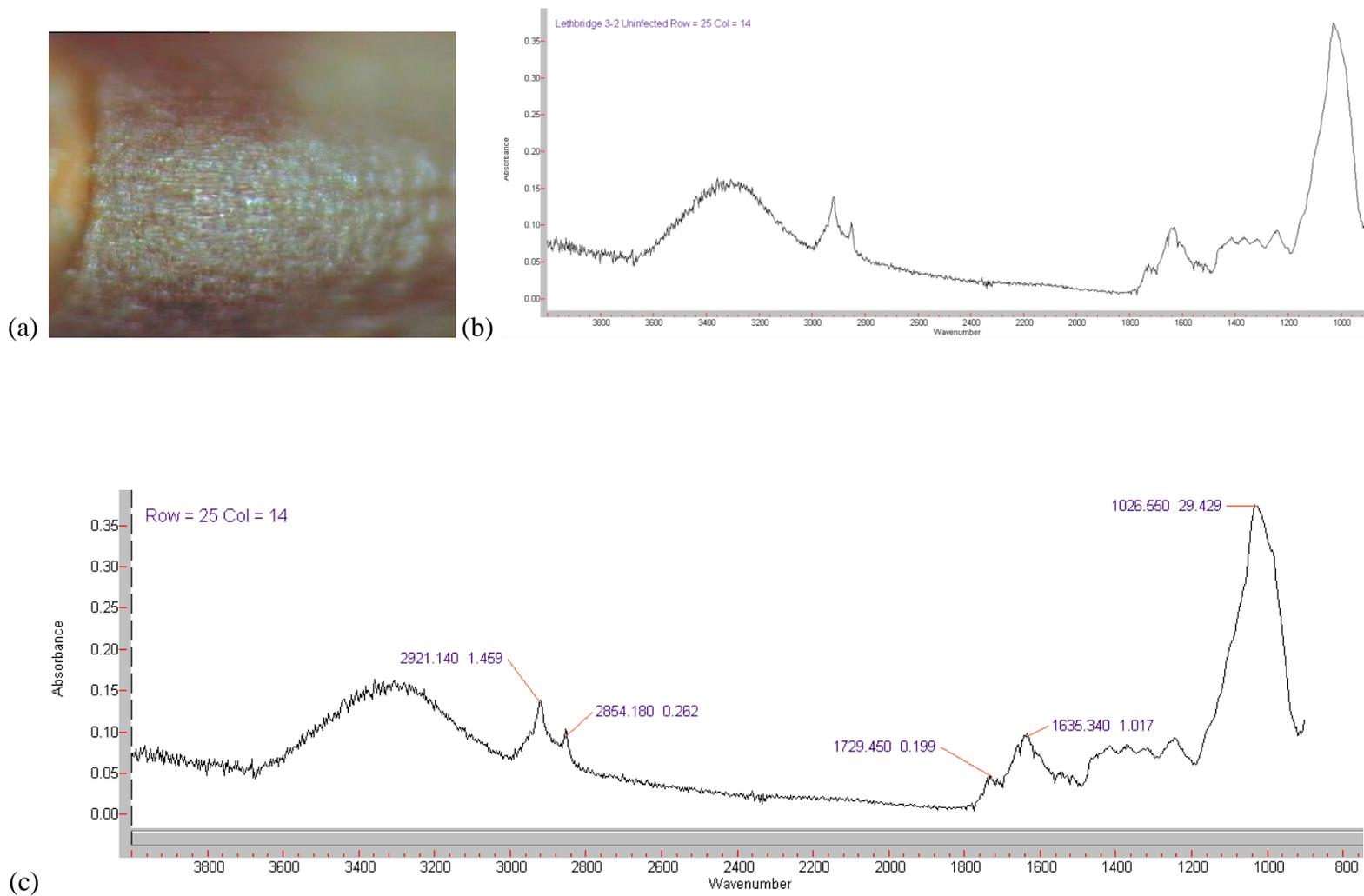
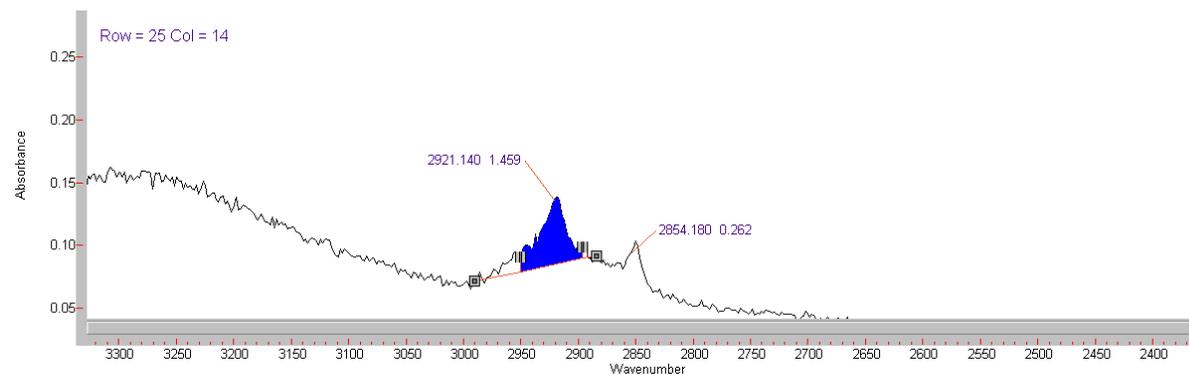
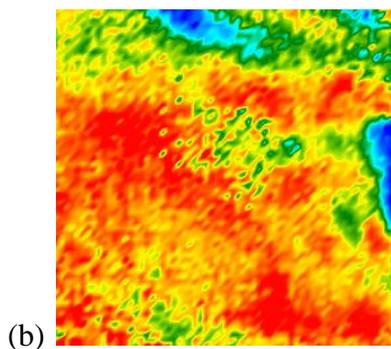
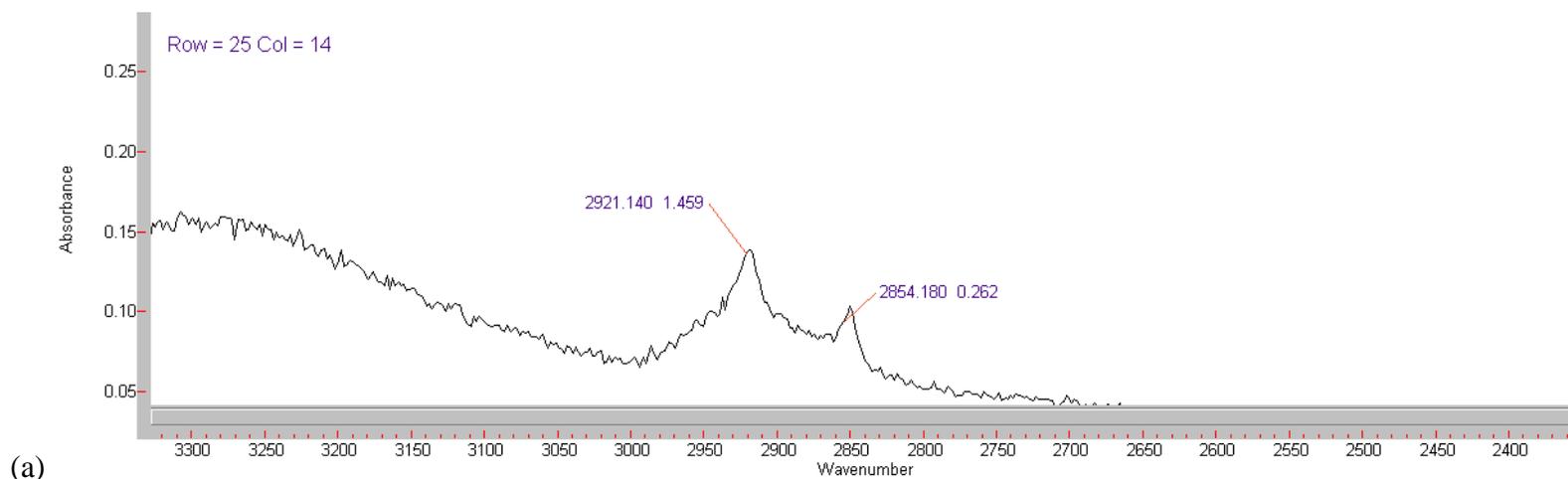
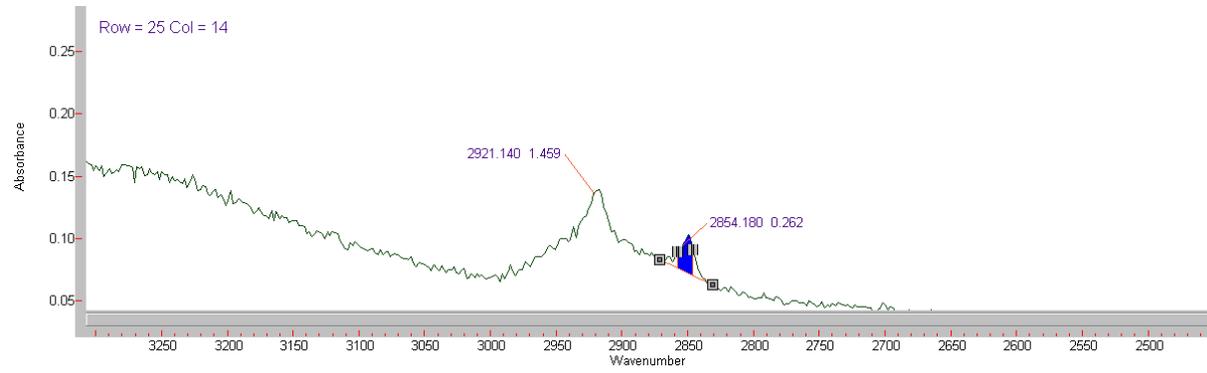
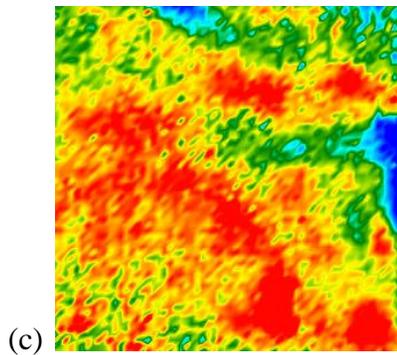
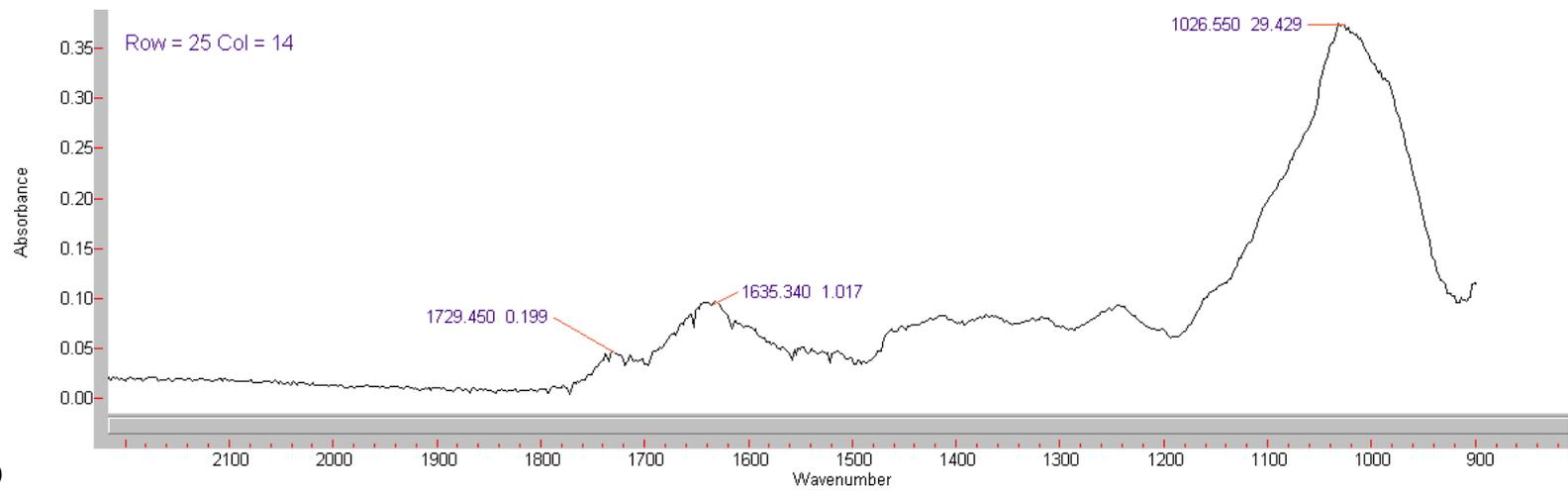


Figure F.6 Typical spectral functional groups and corresponding frequency slices for the CH₂ asymmetric (b) and CH₂ asymmetric (a) functional groups (a), the lipid carbonyl (e), protein amide one (f), and cellulose (g) functional groups (d) for an uninfected wheat kernel in the brown soil zone





(d)



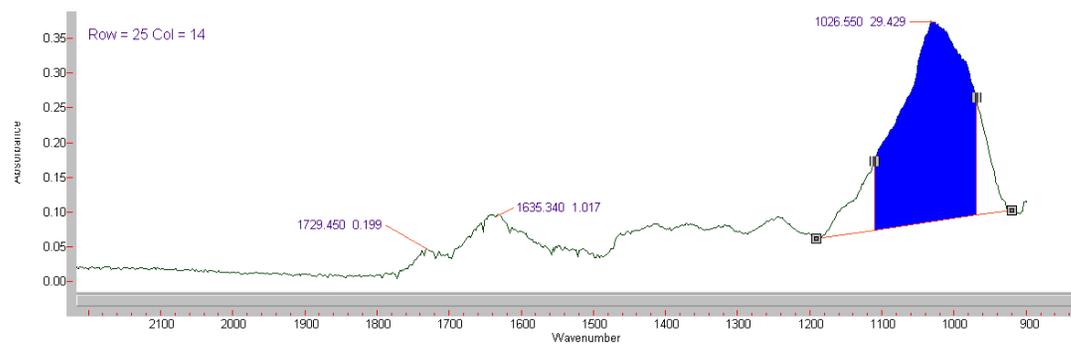
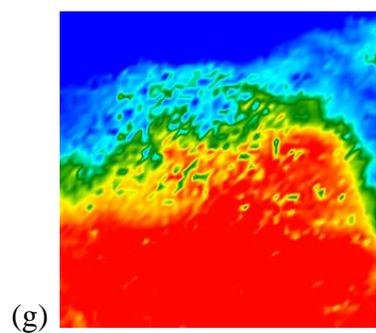
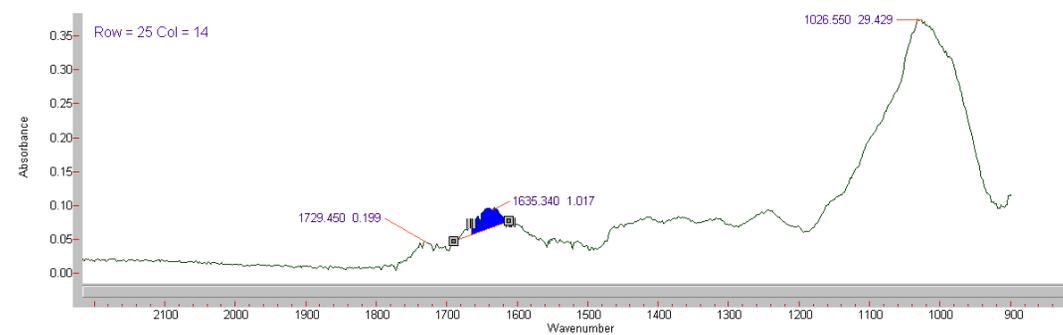
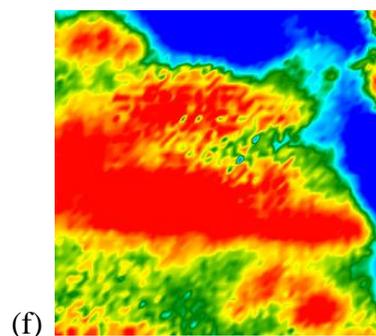
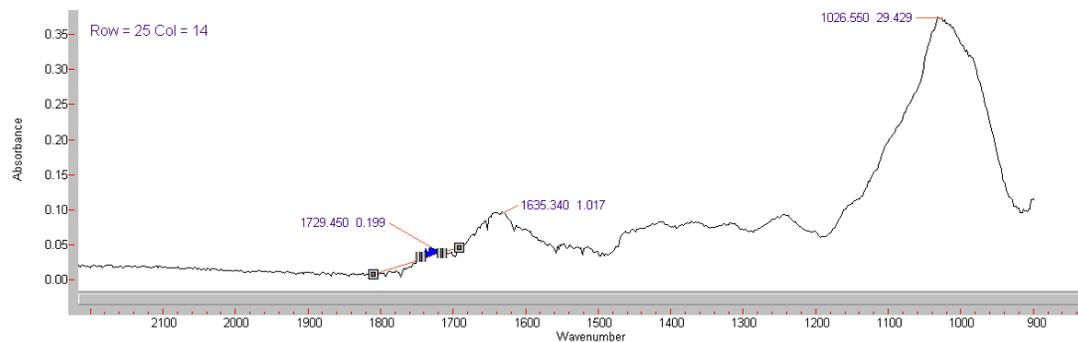
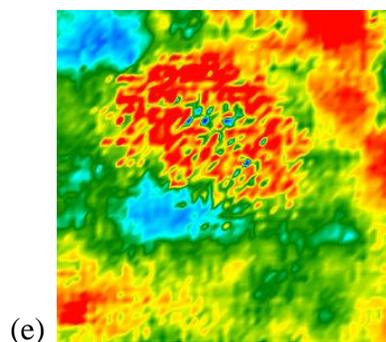


Figure F.7 Typical wheat kernel surface (a), typical spectrum (b), and details of functional groups (c) for an infected wheat kernel in the brown soil zone

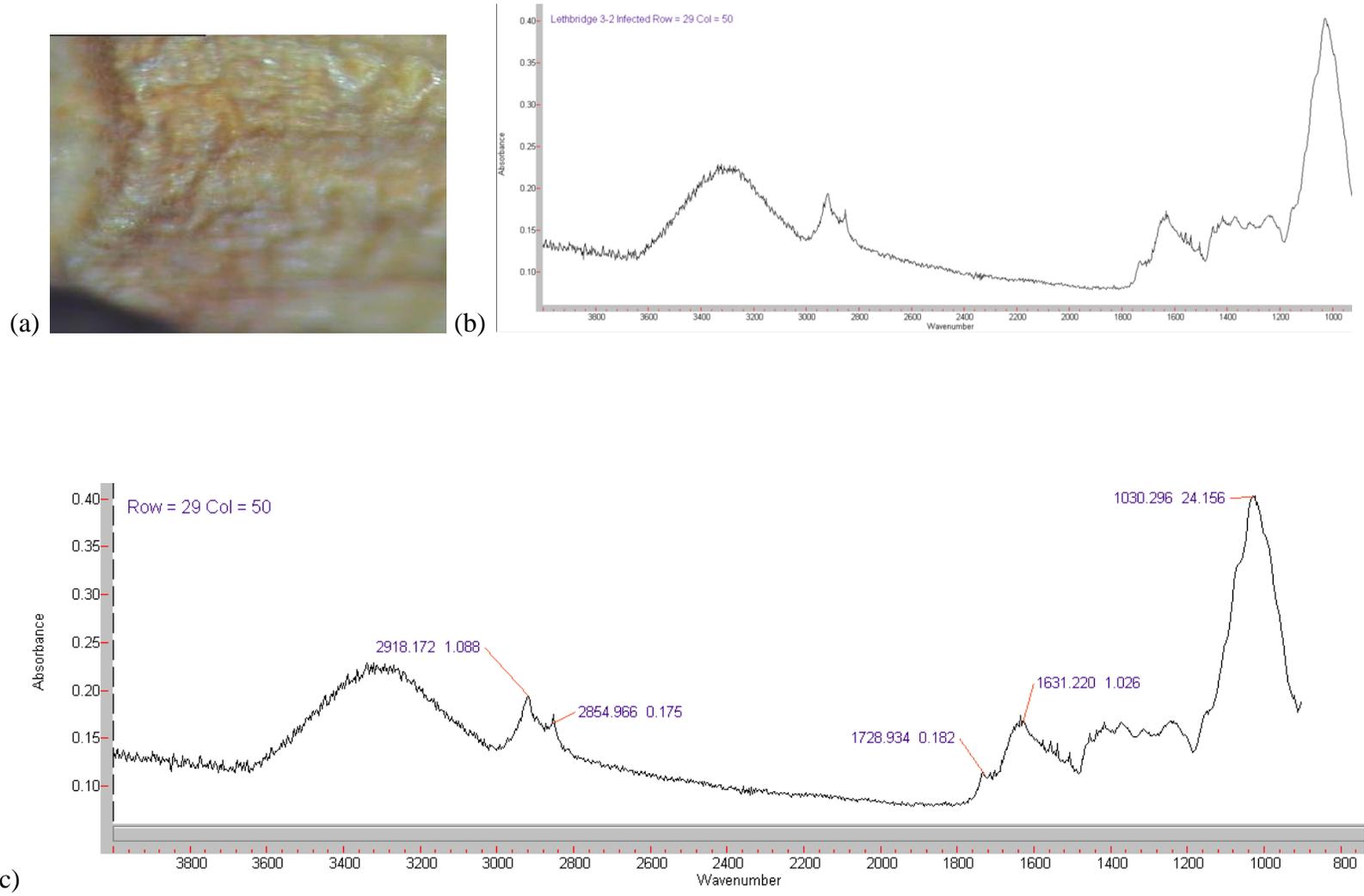
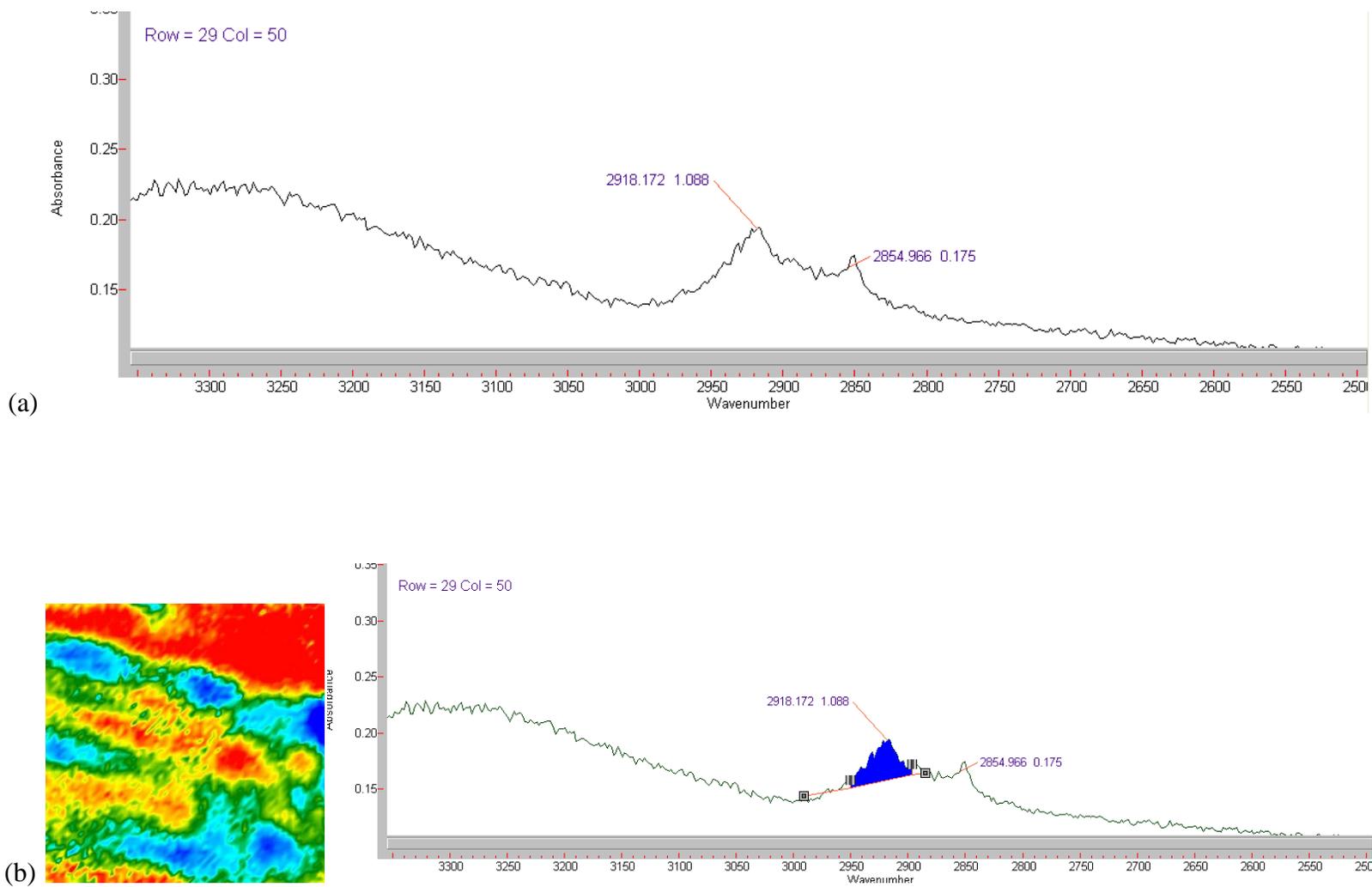
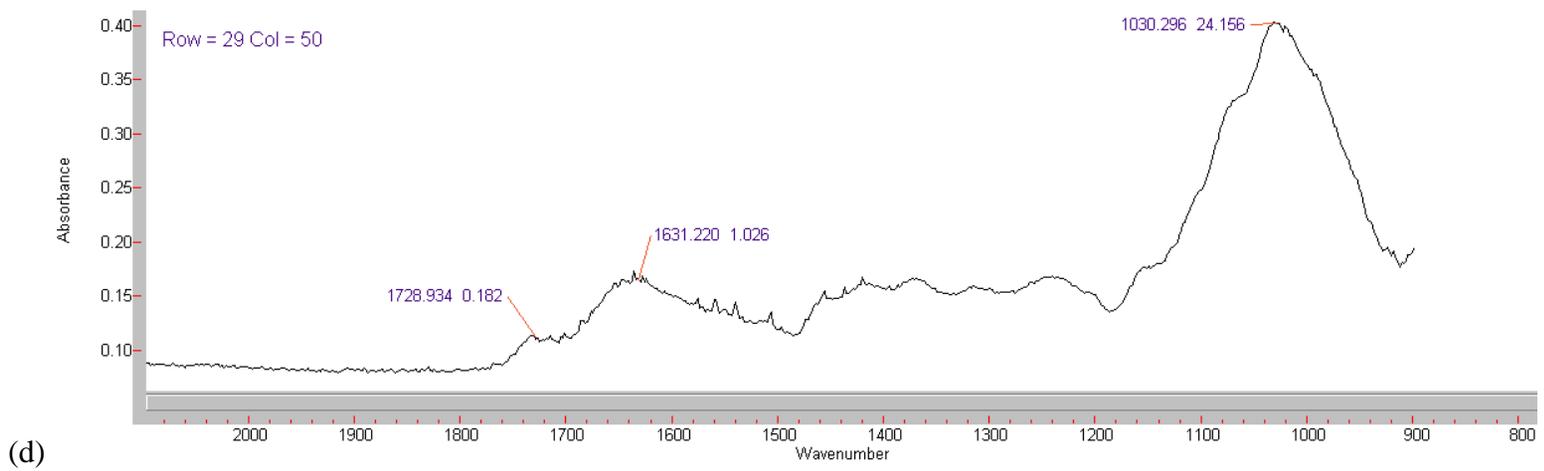
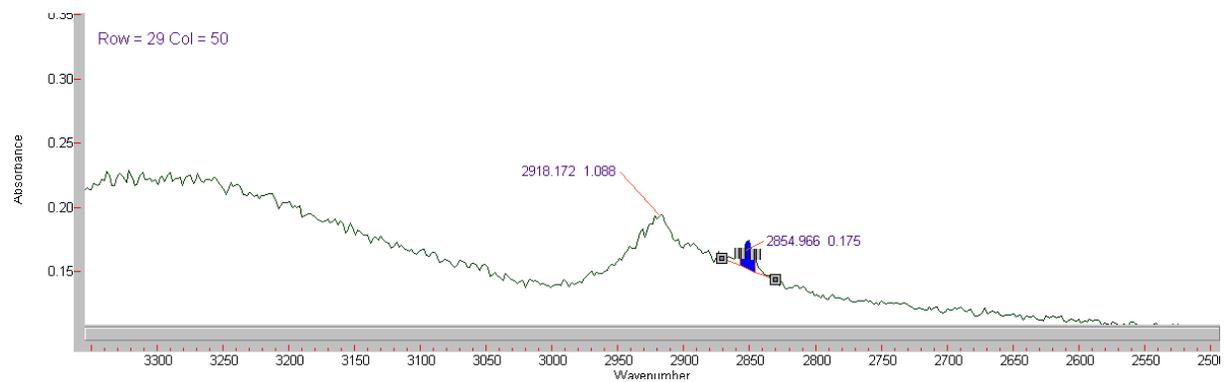
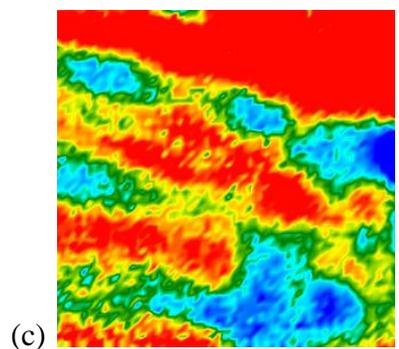


Figure F.8 Typical spectral functional groups and corresponding frequency slices for the CH₂ asymmetric (b) and CH₂ asymmetric (a) functional groups (a), the lipid carbonyl (e), protein amide one (f), and cellulose (g) functional groups (d) for an infected wheat kernel in the brown soil zone





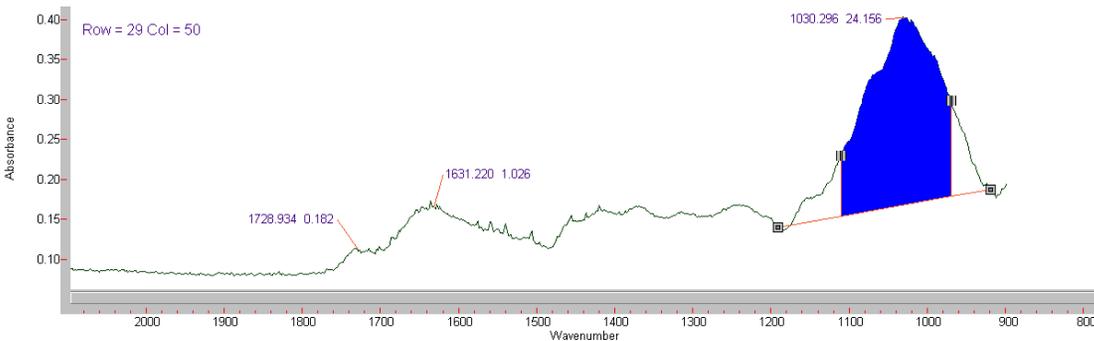
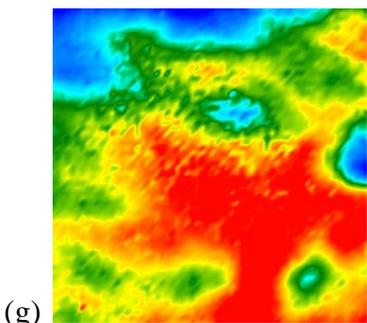
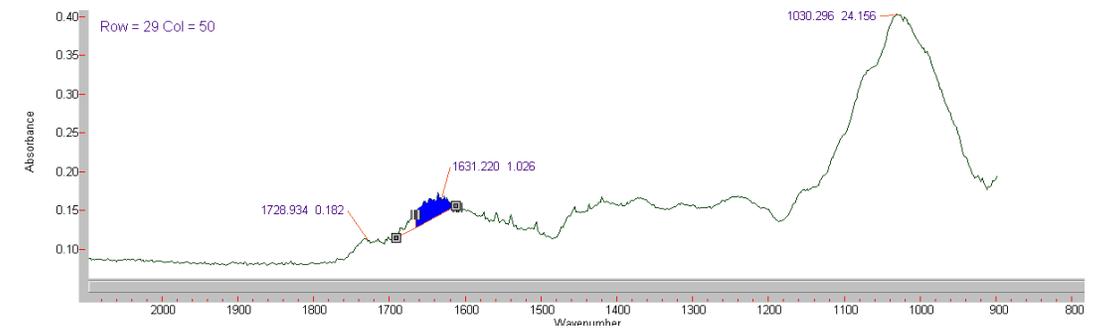
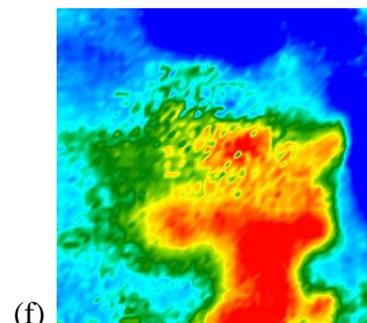
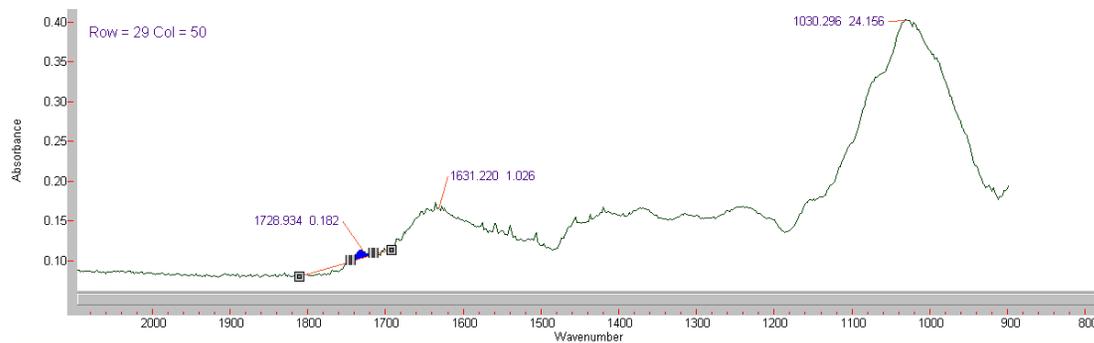
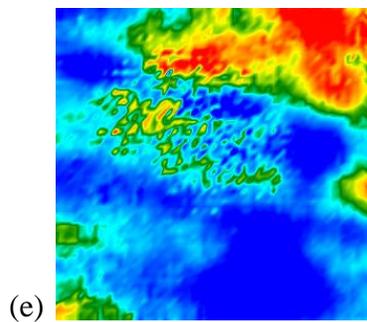


Figure F.9 Comparison of uninfected and infected spectra for kernels in the brown soil zone

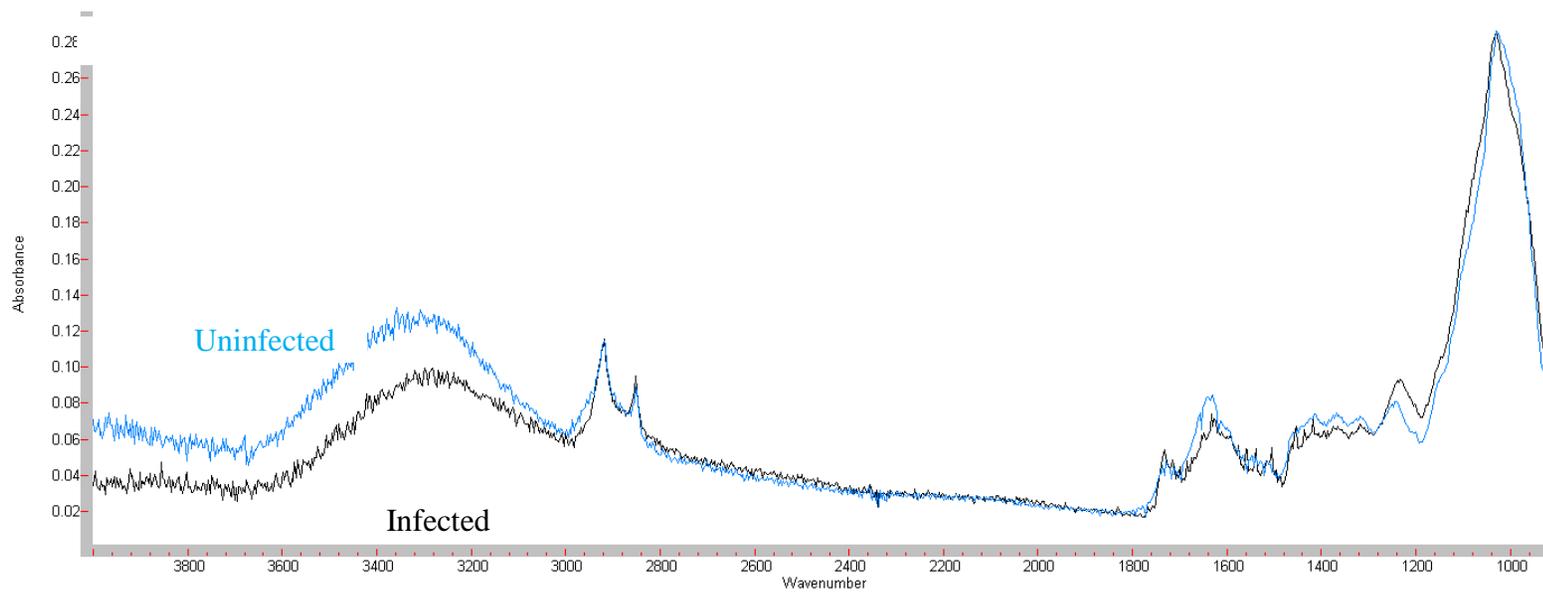


Figure F.10 Typical wheat kernel surface (a), typical spectrum (b), and details of functional groups (c) for an uninfected wheat kernel in the dark brown soil zone

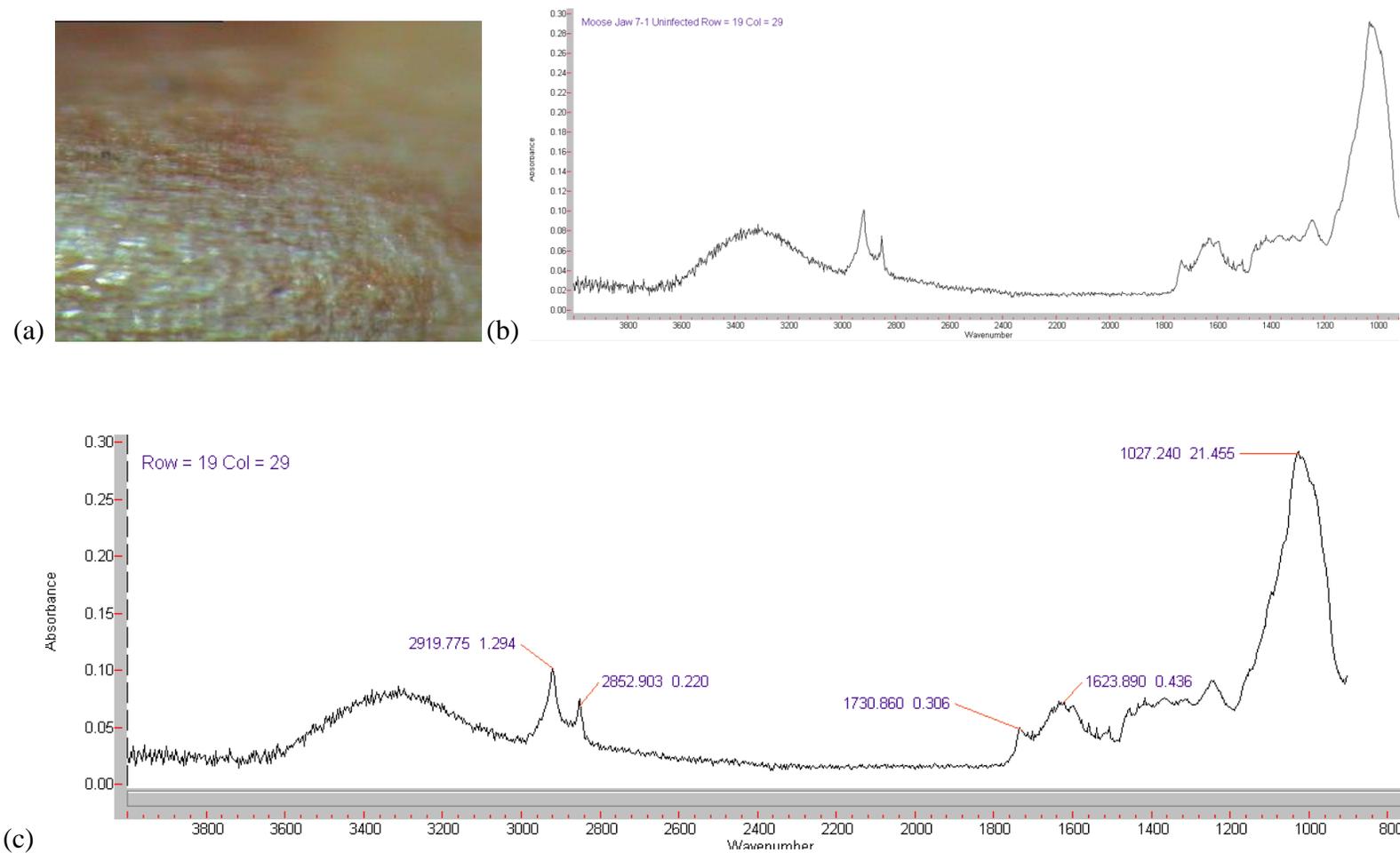
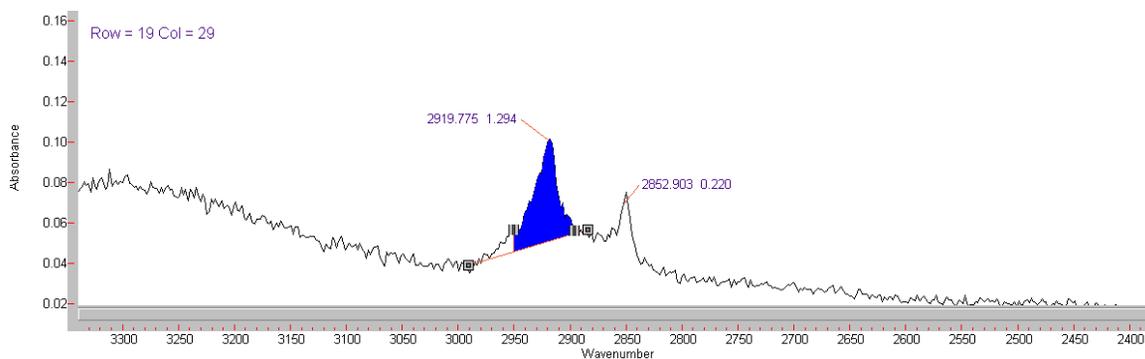
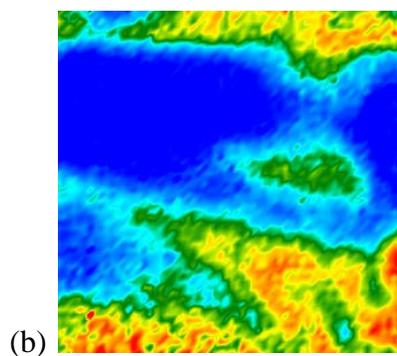
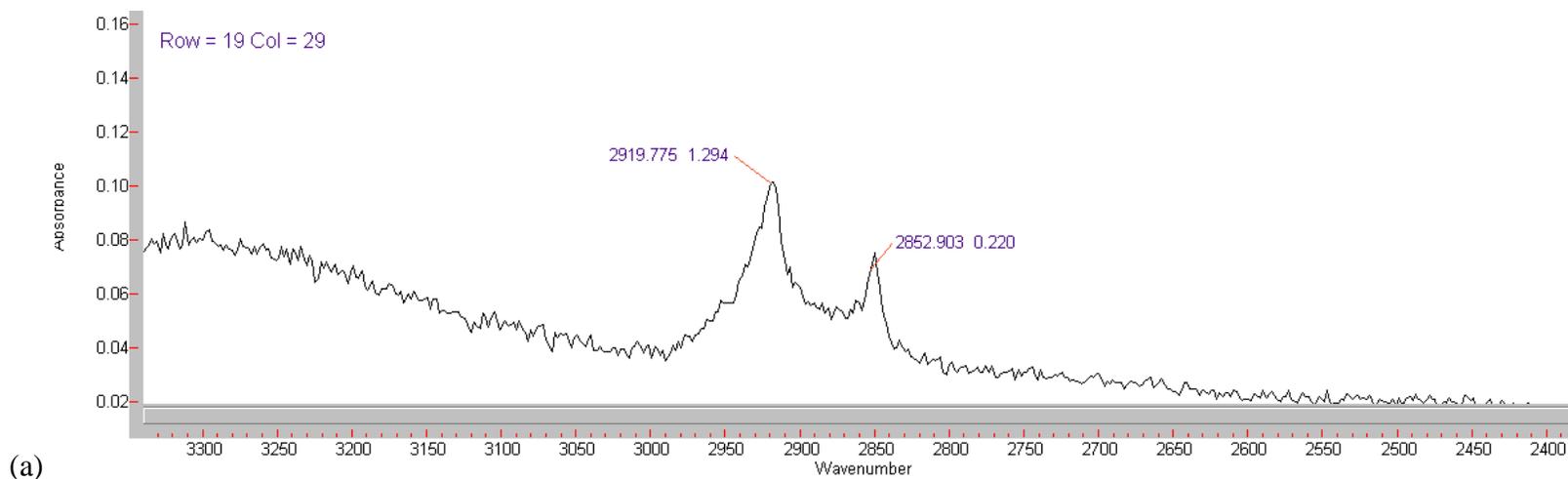
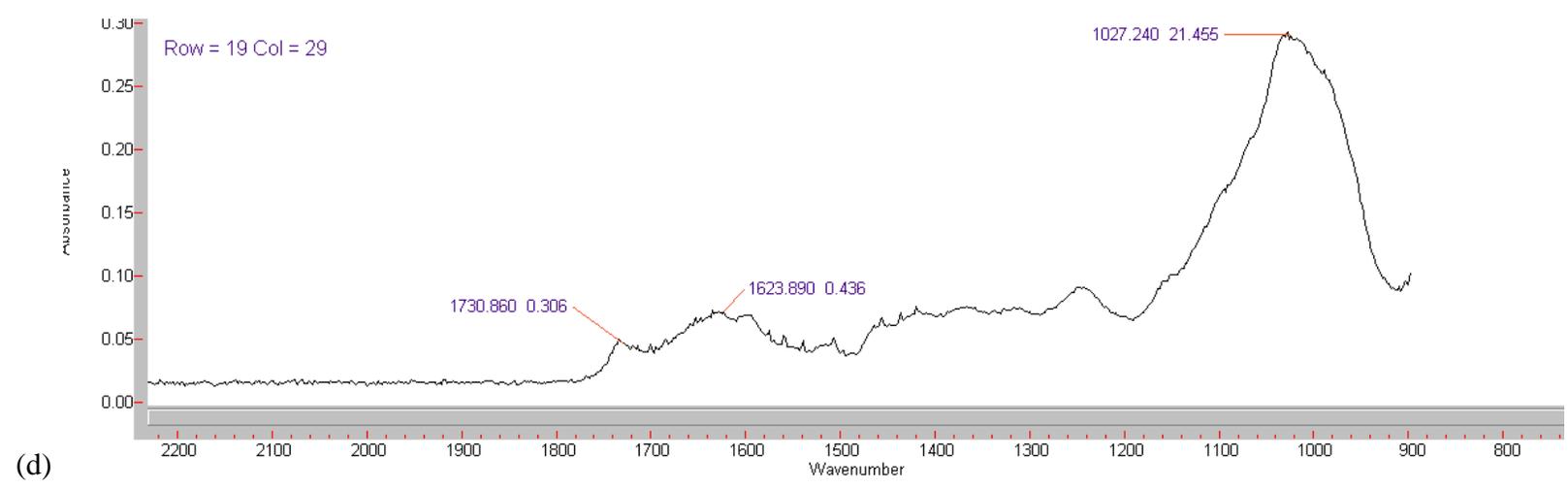
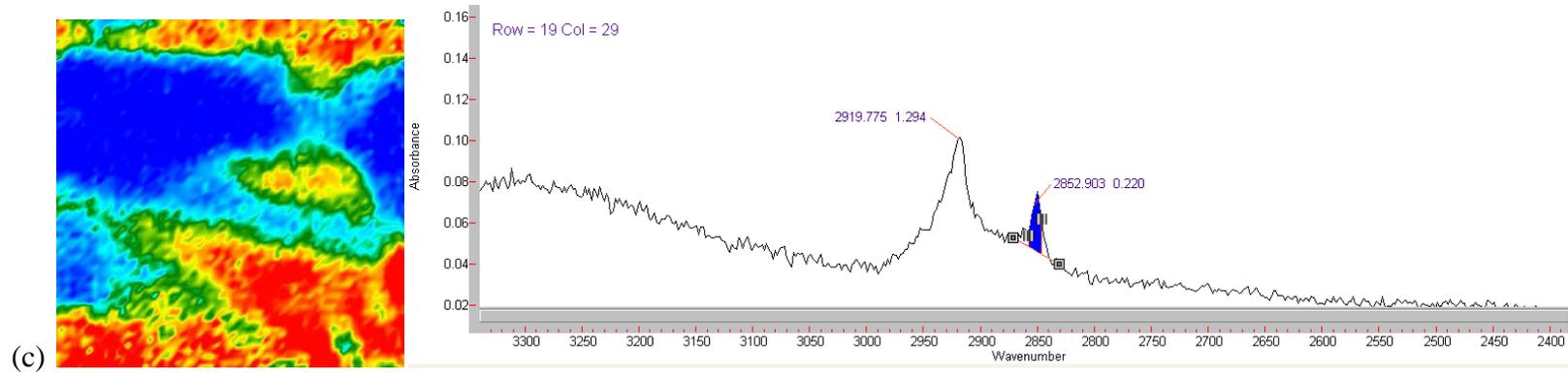


Figure F.11 Typical spectral functional groups and corresponding frequency slices for the CH₂ asymmetric (b) and CH₂ asymmetric (a) functional groups (a), the lipid carbonyl (e), protein amide one (f), and cellulose (g) functional groups (d) for an uninfected wheat kernel in the dark brown soil zone





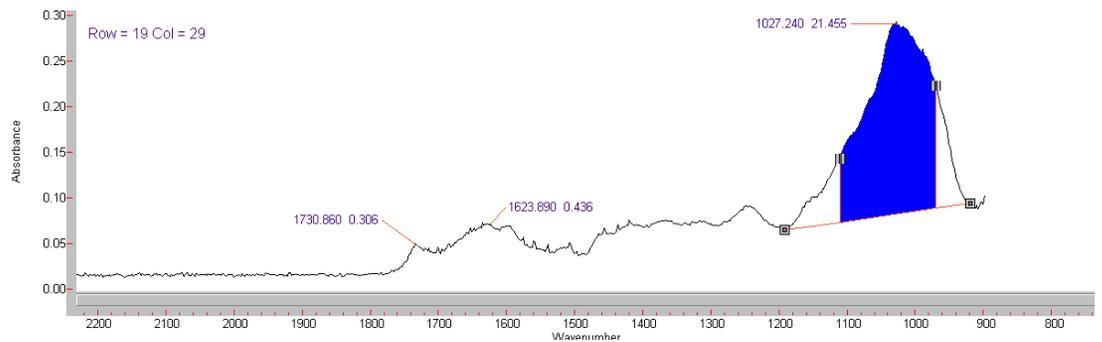
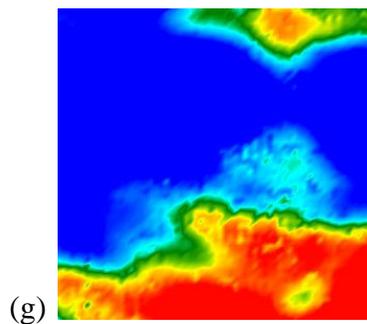
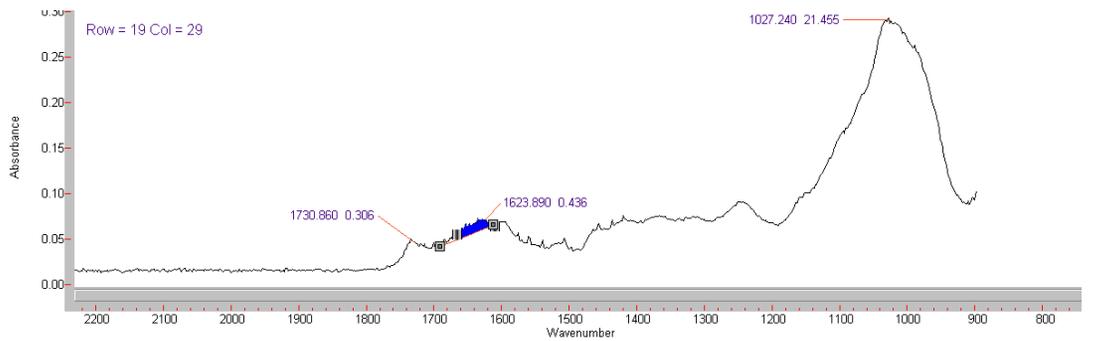
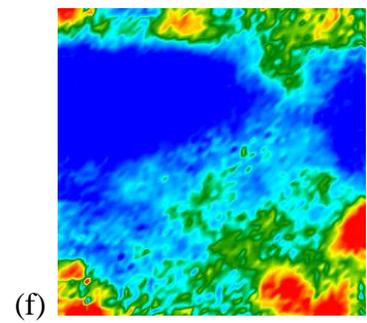
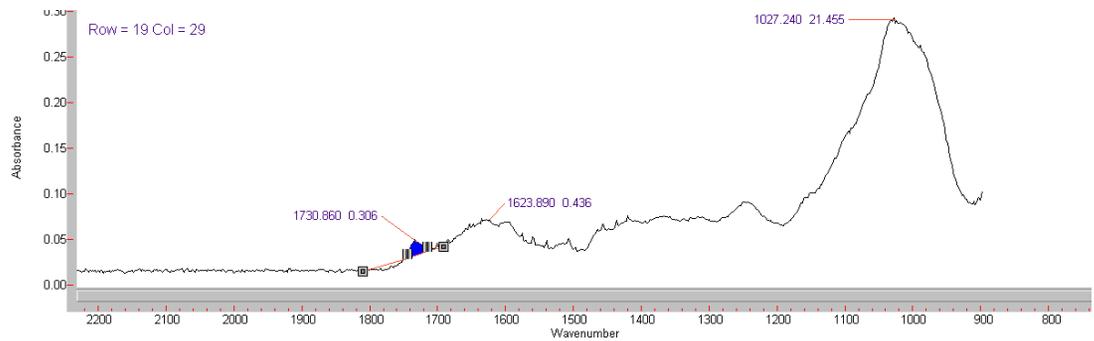
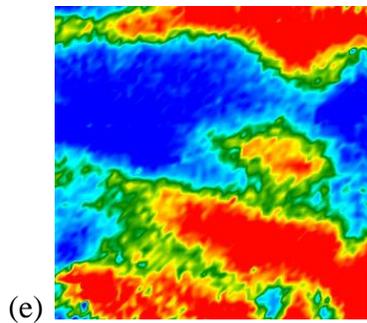


Figure F.12 Typical wheat kernel surface (a), typical spectrum (b), and details of functional groups (c) for an infected wheat kernel in the dark brown soil zone

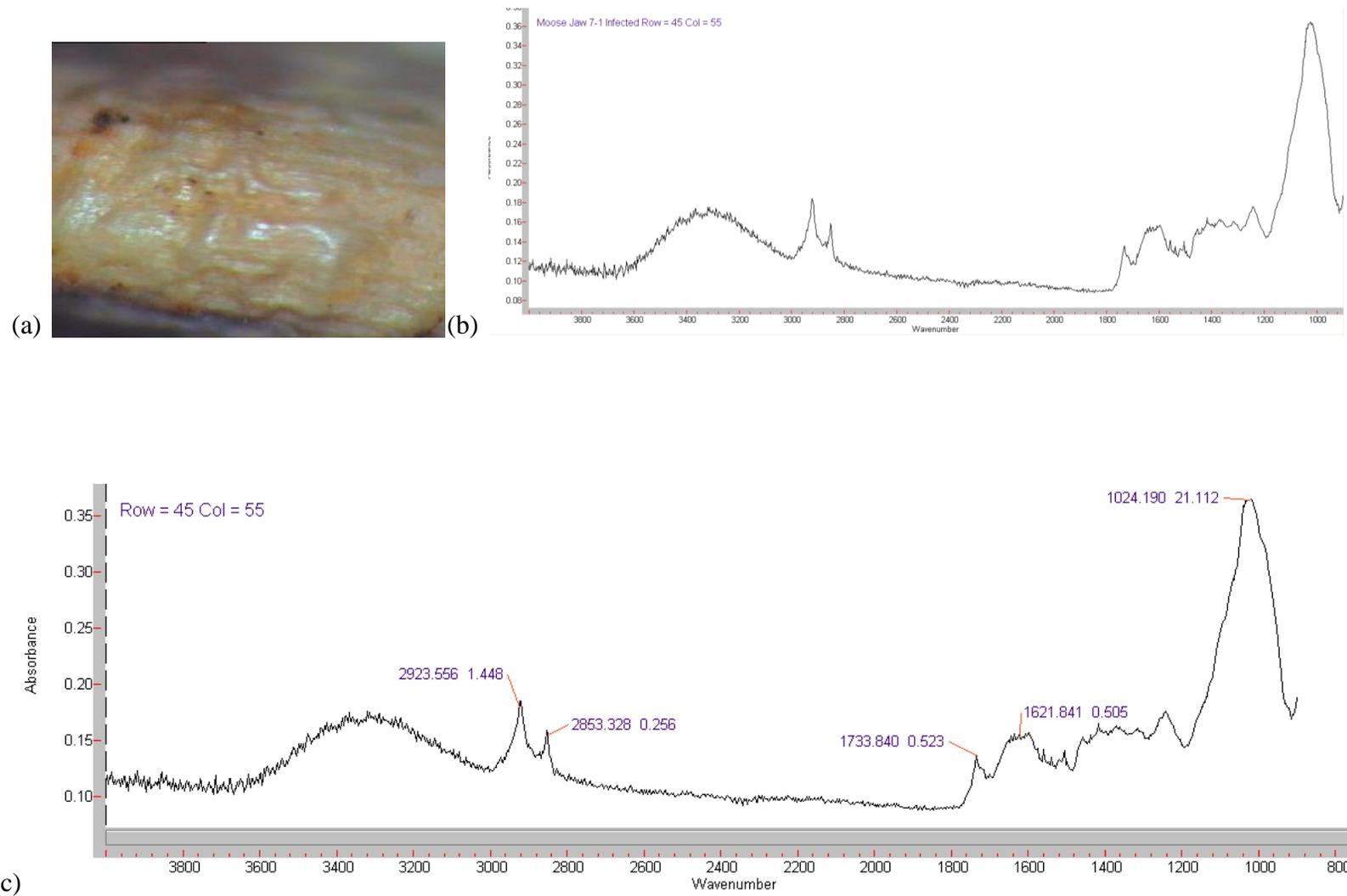
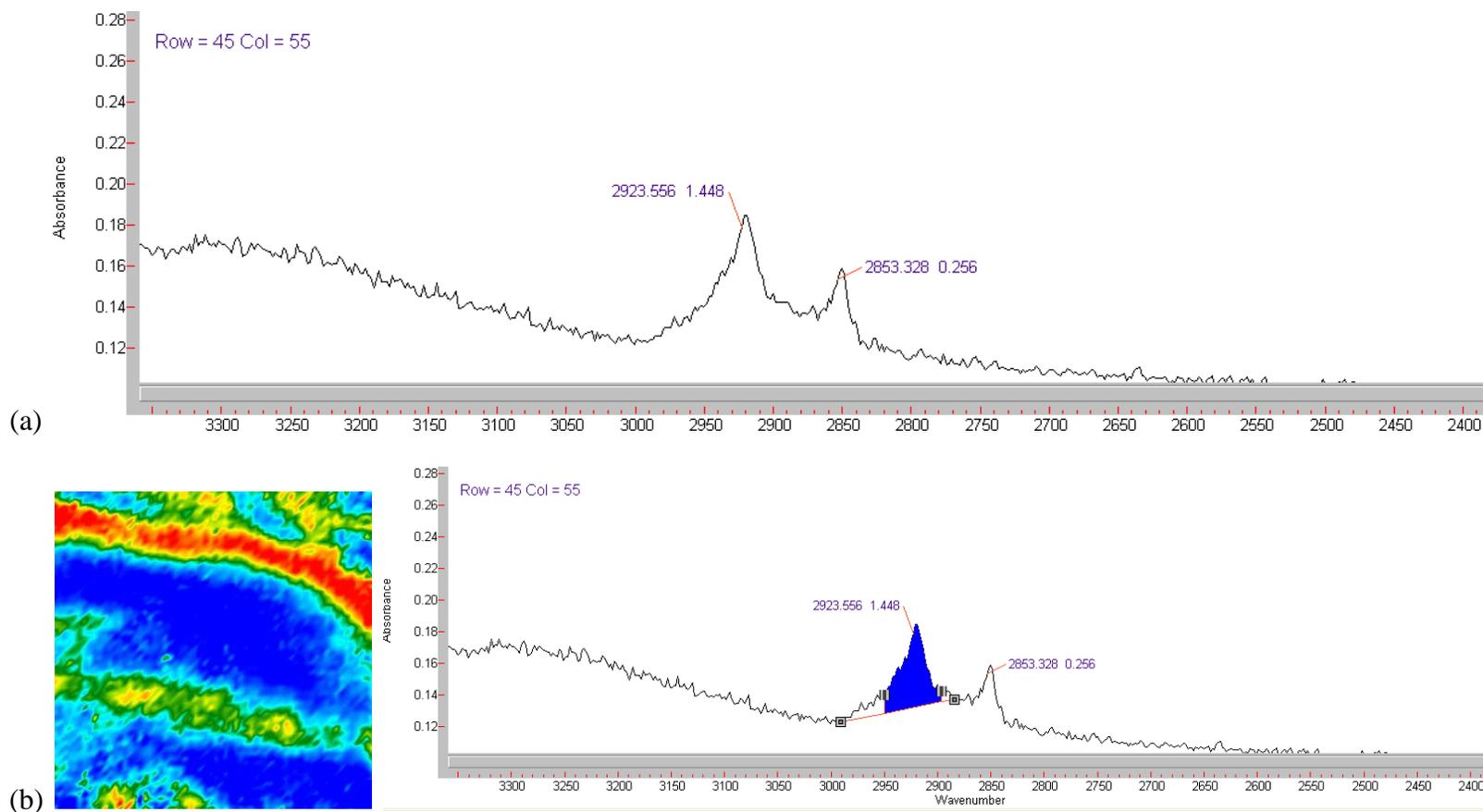
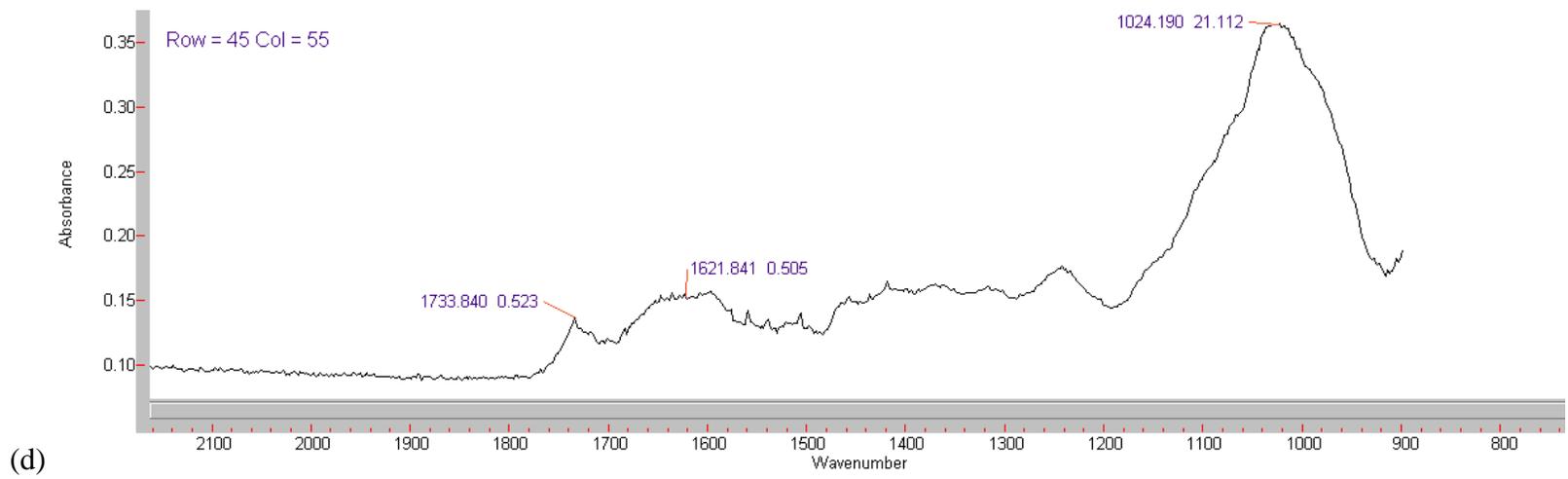
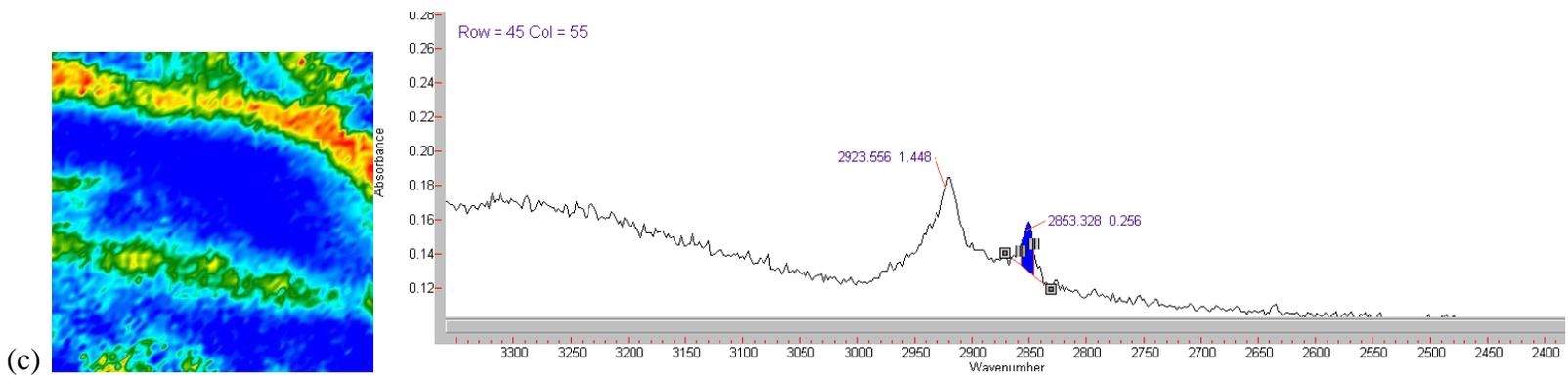


Figure F.13 Typical spectral functional groups and corresponding frequency slices for the CH₂ asymmetric (b) and CH₂ asymmetric (a) functional groups (a), the lipid carbonyl (e), protein amide one (f), and cellulose (g) functional groups (d) for an infected wheat kernel in the dark brown soil zone





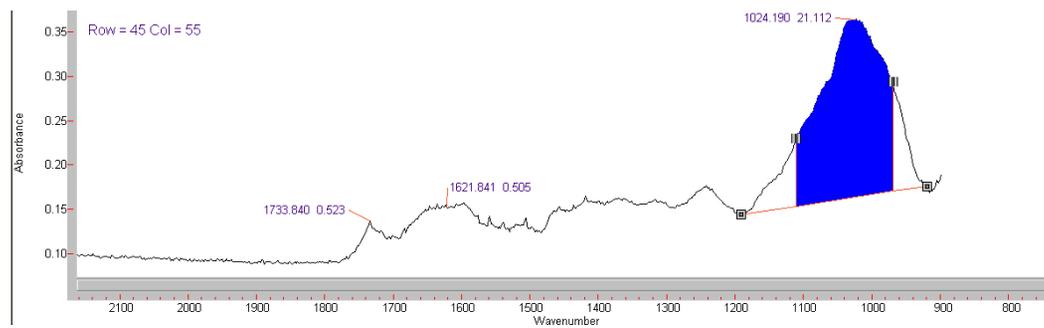
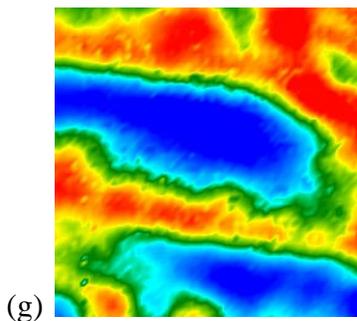
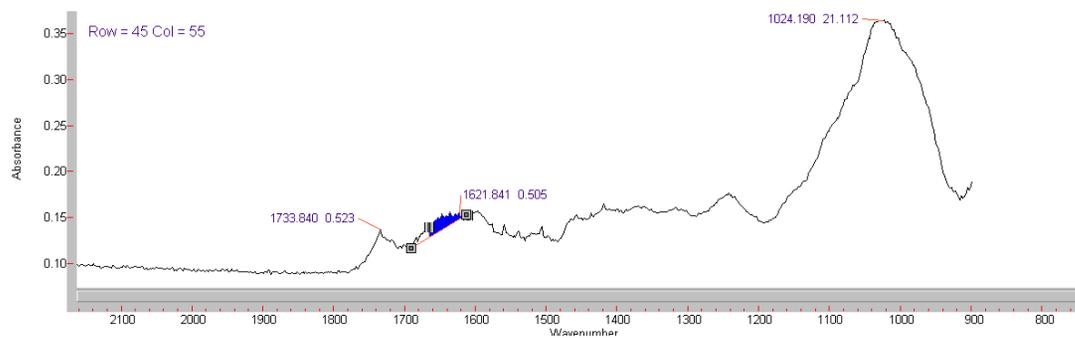
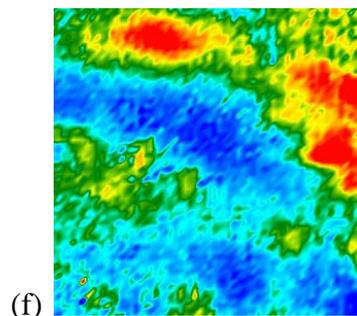
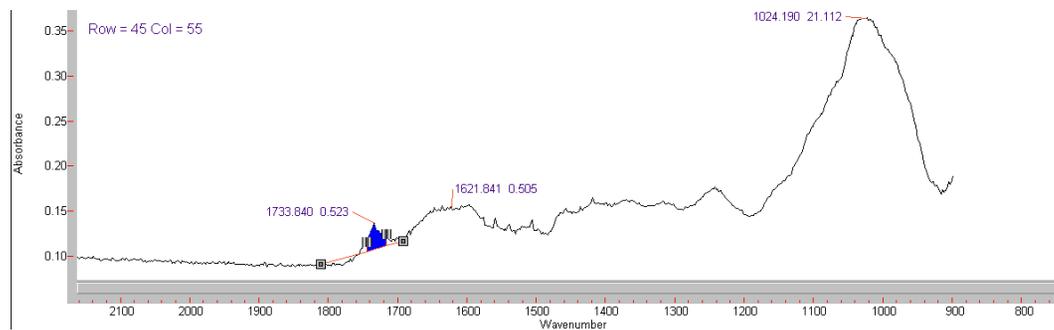
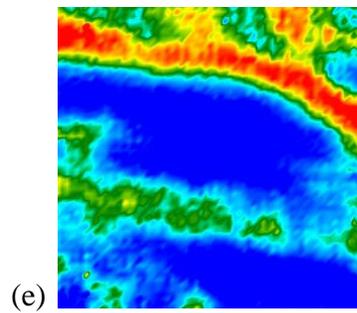


Figure F.14 Comparison of uninfected and infected spectra for kernels in the dark brown soil zone

