CLASSIFICATION OF BULK GRAINS USING THEIR REFLECTANCE CHARACTERISTICS

A Thesis
Submitted to the Faculty of Graduate Studies
The University of Manitoba
in partial fulfillment of the requirements for the degree of

Master of Science

by

Aravind Mohan Lokhamoorthi

Department of Biosystems Engineering
University of Manitoba
Winnipeg, Canada
R3T 5V6

0-612-76791-4

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

*****

COPYRIGHT PERMISSION

Classification of Bulk Grains Using Their Reflectance Characteristics

BY

Aravind Mohan Lokhamoorthi

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirement of the degree

of

MASTER OF SCIENCE

Dedicated to Thirupati Lord Venkateshwara

ABSTRACT

The grain handling system in Canada would benefit from research involving machine vision systems that produce a more consistent, error free, fast and reliable technique for grain grading than that is currently available. It was hypothesized that the machine vision system could be improved by using reflectance characteristics as one of the parameters in classifying grain. The reflectance characteristics of seeds from seven cereals and buckwheat, 10 pulses, three oilseeds and 25 specialty crops were recorded using a spectrophotometer (Model: Cary 5, Varian Canada Inc., Mississauga, ON). The effects of the growing region, seed moisture content, and foreign material content in bulk samples, on the reflectance characteristics of Canada Western Red Spring (CWRS) wheat were also determined.

From the reflectance curves, 465 features based on the ratios, slopes and slope-ratios of the reflectance data were extracted and tested as three models for classification. Procedure STEPDISC was used to rank the features and the top 20 features were used in Procedure DISCRIM for classification. A back Propagation Neural Network (BPNN) was used to collect the weights of the individual features and the top twenty features were used to test the classification accuracy. Ratio features and the slope-ratio features were more successful in classifying than the slope features.

BPNN and discriminant analysis performed similarly in classifying bulk grain. The top twenty features consisted of features from many regions of the electromagnetic spectrum. These classifiers were not successful in classifying the

effects of the growing regions and crop-year, moisture content or foreign material content of CWRS wheat, i.e. these parameters do not affect the reflectance characteristics significantly.

ACKNOWLEDGEMENT

I wish to express my gratitude towards Dr. D.S. Jayas for his guidance, encouragement, and immense support throughout the period of my study.

I am also grateful to Dr. N.D.G. White for his valuable suggestions. I am also thankful to Dr. Lakhdar Lamari for serving on my advisory committee.

I thank the Natural Sciences and Engineering Research Council of Canada for providing the financial support.

I also thank Brett Young Seeds and Agriculture and Agri-Food Canada for providing the seeds.

I am also grateful to my wife Vandana and daughter Smrti for their emotional support during the tougher times.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Canada is a world leader in the export of quality grains. Grading has been the main tool in maintaining a highly consistent quality standard and in the success of Canada in the world export market. This has been made possible by constant research and application of the research to the grain industry. There is a potential for improvement in grading procedures by the introduction of an automated machine vision system (Majundar et al. 1996a, 1996b; Shatadel et al. 1994; Karunakaran 2002, Paliwal 2002, Visen 2002). The introduction of such a system would increase the efficiency of handling, storing and shipping of grain.

A machine vision system consists of an image sensor, such as a camera, to acquire the image and a computer to record, analyze and determine the specific characteristics of grain such as morphology, reflectance, texture and color. A comparatively easy feature that could be used in a real-time industrial application would be the reflectance characteristics, at selected wavelengths, which can be rapidly and easily measured using filters. The reflectance characteristics have been studied since the 1970s and have been used to determine the protein levels in wheat (Hawk et al. 1970; Panford. 1987). Although reflectance has been used for a long time, not much emphasis has been given to nondestructive testing of samples and to the assessment of reflectance characteristics for their ability to classify grain using the ultraviolet, visible and the near-infrared spectrum.

To provide a fast and nondestructive classification technique for bulk samples, the objectives of the study were: (i) to determine the reflectance characteristics of various bulk grain samples in the 320 to 1880 nm spectrum using a spectrophotometer;

(ii) to determine the potential of the reflectance data in classifying various bulk samples; and (iii) to determine the effects of the growing regions and crop-year, the seed moisture content, and the amount of foreign material present in a sample of Canadian Western Red Spring (CWRS) wheat on the reflectance characteristics.

2: LITERATURE REVIEW

2.1 Canadian Grain Grading

A.W. Wood defines grading as "the segregation of heterogeneous material into a series of grades reflecting different quality characteristics of significance to users" (Cited by: Canada Grains Council 1982). Grading the grain helps make a transaction without the presence of the grain. It is also the means of increasing net financial returns to the producer and value to the consumer.

Canada produces, on average, about 55 Mt of grains annually, of which 60% is exported (CIGI 1993). To compete with major grain exporters such as the United States and Australia, it is necessary for the Canadian Grain Commission to set quality standards with stringent tolerance limits to maintain Canada's share of the export market. The Canadian grain grading system is respected globally and other countries follow some of the aspects set by this system. The Canadian Grain Commission is the sole regulatory agency responsible for assigning a numerical grade to the grain and also for monitoring the grain quality as it moves through the commercial handling system consisting of the farmer, primary elevator, secondary elevator, transfer elevator, terminal elevator, rail or ship and the consumer.

The Canadian Grain Commission takes into account a number of characteristics when grading the grain, namely the test weight, vitreousness, moisture content and foreign material content which are measured objectively and variety and soundness which are measured subjectively. The grain when delivered by the farmer to the primary elevator is inspected thoroughly and given a grade by the elevator personnel. Canadian Grain Commission inspectors once again grade the grain, when it arrives at the transfer or terminal elevators from the primary elevator.

The grading is facilitated by the standard measuring systems and standard samples prepared for reference every year. Although the inspectors who grade the grains are highly trained, the possibility of erring is there as a result of inconsistency, fatigue and other internal and external factors (Kohler 1991). This drawback can be alleviated with a machine vision system, which is a technological advancement in the grading technology dealing potentially with all the drawbacks in the current system.

2.2 Machine Vision System

A machine vision system is a combination of a camera or a scanner, which acts as the eye of the system and a computer, which acts as the brain to replace the human involvement in the grain grading process. This system, installed at elevators, could act as an inspector and/or assist the grain inspectors in the grading of grain. This has led to an interest in the grain industry to use online monitoring systems. The fast and accurate evaluation of the contents of a sample of grain by this system could be used to obtain optimum cleaning strategies for the grain or for a completely automated system in some of the elevators during the unloading of grain from rail cars to make the appropriate segregation decisions (Shatadal et al. 1995).

The eye of the system (camera) captures the image of the sample. An appropriate converter is used to transform the image to the required format and is then passed on to the computer for analysis. The computer in turn records the image and using custom made programs extracts features based on color, texture and morphology of objects to facilitate the classification, sorting, or grading process.

Shatadal et al. (1995) described a machine vision system that was tested for differentiating hard red spring (HRS) wheat or barley from six types of large seeds and five types of small seeds and differentiating small and large seeds using morphological features. They reported that classification was more than 99 % accurate for HRS wheat and barley from all other seeds types and there was a very large misclassification when classifying the mixed small seeds or the mixed large seeds themselves.

Morphological features were used in classifying nine cultivars of milling quality wheat from eastern Ontario, consisting of hard red winter (HRW) wheat, hard red spring (HRS) wheat, and soft white winter (SWW) wheat (Symons and Fulcher 1988a). The classification accuracy of SWW wheat was 100 % while for HRW wheat and HRS wheat it was approximately 80 %. A second experiment was carried out to test the potential of discrimination within cultivars and to test the effect of environment (Symons and Fulcher 1988b). The results showed that a classification accuracy of more than 80 % could be achieved and there was an influence of the growing conditions, which affected the kernel morphology, subsequently influencing classification.

In another study, the application of digital imaging for the classification of wheat cultivars according to the kernel type was studied (Neuman et al. 1986). A perfect, 100 % classification accuracy was obtained for four durum wheat types and Canada Western Red Spring wheat and varying classification accuracies ranging from a high 96 % and a low 15 % was obtained for Canada Western Red Winter, Canada Western Soft White Spring, Canada Utility and Canada Prairie Spring wheat.

A computer-controlled laser scanning system was developed to acquire three-dimensional images of cereal grains. A separate image was acquired to represent the light reflected from the kernel surface and a combination of the images was used to acquire the features, which could classify 92 – 94 % of soft white winter and Tyee (club) wheat (Thompson and Pomeranz 1991). A similar image analysis system could also discriminate cereal grains and weed seeds and between soft white and club wheat, two and six rowed barleys, and rye and triticale kernels, with a classification accuracy of 99.5 % between wheat and nonwheat, 90 % between soft white and two club wheats and two and six rowed barley, and 90 % between rye, triticale and other grains (Chen et al. 1989). The probability of touching grain kernels is very high when the grain sample is presented for image acquisition. Shatadal et al. (1994) developed a disconnect algorithm based on mathematical morphology and tested it for HRS wheat, durum wheat, barley, oats and rye and obtained a mean classification accuracy of 93.3 %.

Machine vision methods have also been used for sorting of stone fruits such as peaches, nectarines and plums on the basis of their shape and surface defects (Singh and Delwiche 1994). Uniformly illuminated samples were imaged in the visible and near-infra-red spectra and were processed to detect major defects like cuts, bruises, scars, wormholes and misshapen fruits. An overall classification accuracy of 71.4% was achieved and the errors were primarily due to the natural variability in the features.

A machine vision system using features consisting of color, shape of stem cut and cap veil opening was used to evaluate the grading of mushrooms and compared to two trained inspectors (Heinemann et al. 1994). The machine vision system gave a misclassification of 8 to 56 % with an average of 20%. A disagreement of 14 to 36 % was reported between the two inspectors. They concluded that the machine vision system was better in grading consistently than human inspectors who had a bias and lack of consistency.

The machine vision system was used for distinguishing good and greened potatoes and yellow and green 'Golden Delicious' apples with 90 % accuracy (Toa et al. 1995). Machine vision has also been used for inspection and grading of apples, carrots, bell peppers, peaches and soya beans (Reyer et al. 1996)

Machine vision has also been extended in the area of soft x-rays where the images acquired using soft x-rays help in differentiating the insect-infested from the uninfested cereal kernels with a classification accuracy of 93.8 % as opposed to the standard Berlese funnel method, which had a extraction effeciency of 81 % (Karunakaran 2002).

Apart from laboratory experiments, measurements based on the reflectance characteristics have not been used widely (Casady et al. 1993; Howarth et al. 1990; Majumdar et al. 1996b; Hawk et al. 1970). The spectroscopy and reflectance measurements provide: (i) the ability to process and classify bulk grain within a very short period of time; (ii) ease in preparing the test samples; and (iii) low cost of the equipment compared with a machine vision based system using single kernels. These need further research to develop applications for the grain industry.

2.3 Reflectance Characteristics

Use of reflectance characteristics was limited to analyzing the mineral content in rocks, chemical composition of food material and checking the structural integrity of rocks (Eu 1997). Near infra red spectroscopy is a procedure that rapidly detects and measures the chemical composition of biological material and classifies defects, e.g. internal insects in wheat (Dowell et al. 1998).

When white light is passed through a material it could be selectively absorbed, rendering the resulting emergent radiation as wavelength dependant, which is perceived as color. Color likewise is imparted by selective scatter of radiation like the blue color of the sky imparted as a result of scatter of sun light by atmospheric particles. If the concentration of the scatter centers were very large then a considerable portion of the incident light would be returned to the surface. It would then be possible to describe the material in terms of its reflectance characteristics ranging from an ideal mirror–type surface that reflects light in one angle only, to an ideal matte surface, which reflects uniformly in all angles, which is usually called diffuse reflectance (Williams and Norris 1987).

The most common method of measurement to obtain reflectance is done with a reference to a standard material of known characteristics. A single-beam mode is used where the instrument is calibrated to read 100 % reflectance for the standard material and percentage reflectance measured later with the sample in place of the standard material.

Reflectance characteristics can identify carrots for size, shape and four defects, namely dry rot, soft rot, black crown and cavity. Except for the cavity defects, the reflectance characteristics could differentiate between all other specifications based on reflectance data in the range of 535 and 722 nm. The misclassification of cavity spots was due to the small size of the cavity spots as the exposed samples had both normal and defective areas (Howarth et al. 1990).

Using a spectrocolorimeter, the percent reflectance over the visible spectrum (400 – 700 nm) was used as a rapid first stage identification to identify bulk samples of cereals (namely hard red spring wheat, durum wheat, Canada prairie spring wheat, 6-row barley, feed barley, oats and rye), pulses (white pea bean, pinto bean, black bean, field pea green seeded, dark green speckled lentils, eston lentils and laird lentils), and oilseeds and specialty crops (yellow, oriental and brown mustard, sunflower, flaxseed, canola, and buckwheat). Most of the grains were correctly classified using the reflectance characteristics at wavelengths ranging from 450 to 670 nm but no one wavelength could be segregated to give the required classification (Majumdar et al. 1996b).

Reflectance characteristics with percent reflectance over a wavelength range of 350 - 1800 nm were used to classify eight cereals, three oilseeds, eight pulses, and 27 specialty seeds. Canada western red spring (CWRS) wheat samples at five different moisture contents, with five different foreign material contents, three grades and from 20 growing regions in western Canada were also classified. Thirteen randomly selected features were extracted from the reflectance data and a classification accuracy of 100 % was obtained for the three oilseeds, seven of the eight classes of cereals, five of the eight classes of pulses and twenty of the twenty-seven classes of specialty seeds. Classification of CWRS wheat was not accurate for different grades, samples with different moisture contents, and foreign materials, and from different growing regions. The three best wavelengths in the electro-magnetic spectrum that classified the above-mentioned seeds were 800, 1050 and 1250 nm (Eu 1997).

The reflectance characteristics in the near infrared spectrum were used for a quick analysis of wheat, barley, oats and soybeans for oil, moisture content, and protein. The samples were ground before testing and a standard error of ± 0.22 % for protein content, and ± 0.16 % for moisture was achieved with hard red spring wheat, and 1 – 5 % coefficient of variation with other cereals, oilseeds and legumes. The error was reported to be due to the variation in the method of grinding, which influenced the sample, presented to the sensing unit (Williams 1975).

Using the reflectance characteristics in the infrared region, whole grains of corn and sorghum were tested to determine the moisture contents ranging from 13 – 55 % wet basis, using three wavelengths namely 1.94 μm (a water absorption band), 2.19 μm (a region adjacent to the protein absorption band), and 2.33 μm (a region where starch and oil absorb radiation). Results were compared with electric moisture meters and the standard oven method. The infrared method of measuring the moisture content of whole grain was consistent with the oven method and more accurate, especially in the upper range of moisture content where electric moisture meters were unreliable (Stermer et al. 1977).

Preliminary analysis of the reflectance characteristics were determined for hard red winter, soft red winter, soft red spring, white and drum wheats, white oats, barley, rye, yellow grain sorghum, yellow soy beans, yellow corn and flax (Hawk et al. 1970). The UV region of the spectrum showed no differences within the reflectance of grains and very minimal differences were shown in the near infrared region. Very prominent differences occurred in the visible region for classification between hard winter wheat, oats and grain sorghum. It was concluded that grains could be classified by using grain samples with average reflectance greater or less than the primary grain. It was also shown that by combining any two wavelengths, an admixture (grain other than the primary grain) could be distinguished from the primary grain (Hawk et al. 1970).

Delwiche and Norris (1993), Delwiche et al. (1995) used the reflectance data in the near infrared region to differentiate two classes of wheat, namely, hard red winter and hard red spring. Initially the tests included samples, which were ground. The parameters used were namely NIR-predicted hardness, NIR-predicted protein content, and NIR protein and NIR hardness and the recordings over the entire range of the NIR spectra. The one-parameter models resulted in very poor classification whereas a five-factor principal component analysis resulted in 95 % classification accuracy. The whole grain classification included the first three-year samples in the previous study for calibration and the fourth year samples for verification, using four types of classification algorithms namely, multiple linear regression, principal component analysis with mahalanobis distance, partial least squares analysis, and artificial neural networks. The artificial neural networks produced 95 – 98 % classification accuracy whereas other methods had accuracies varying between 88-95 %.

Bilanski et al. (1984) used spectral reflectance data ranging from 350 to 700 nm from a spectrophotometer on a total of 668 apples consisting of Macintosh, Greening, Spartan, Red Delicious and Northern Spy varieties for bruise detection by dropping each apple from a set height. They reported that a two wavelength derivative model distinguished between good and bruised apples but not a two-wavelength model or a one-wavelength model. The reflectance properties of the apple tissue, however, could not predict the depth of the bruise on the apples. The optimal wavelengths, which produced a consistent result, were in the 552 to 560 nm range.

Upchurch et al. (1990) determined the reflectance characteristics of two bruises of 1 to 4 mm depth on 50 'Red Delicious' apples using diffuse reflectance in the range of 400 to 1000 nm to distinguish bruised and nonbruised areas of unpeeled apples. The two bruises, one on the blush side and the other on the opposite side, were made with the same impact energy using a release pin. It was found that the ratio, normalized difference and derivative models gave the best performance with a total misclassification of 2.5 – 3.5 %, whereas the single wavelength model had the worst classification accuracy misclassifying about 50 % of the apples.

Reyer et al. (1995) performed similar tests on peaches and apricots. In their study, the bruises were created by using a loading head attached to a load cell and the reflectance data were collected using two sensors, the first ranging from 500 – 1000 nm and the second from 700 to 1600 nm. The data were analyzed to select a spectral filter of

750 nm for apricots and 930 nm and 970 nm for peaches which when used in an image acquisition system resulted in a success rate of approximately 65%.

Bittner and Norris (1968) used reflectance data over a range of 250 – 2100 nm to predict maturity of apples, peaches and pears. The three fruits did not show considerable differences in the UV region. All the fruits showed considerable differences in the visible region with an increase in reflectance in the 670 nm region influenced by a decrease in absorption of chlorophyll. Peaches exhibited this attribute more than apples and pears. The red apples had a large decrease in their reflectance properties in the 550 nm region as the maturity progressed and a small increase in reflectance in the 670 nm region in contrast to the Golden Delicious apples, which had a maximum difference at 670 nm as the maturity progressed. Pears showed that one wavelength in the range of 550 – 620 nm would help provide the best index for predicting maturity (Bittner and Norris 1968).

2.4 Spectrophotometer

When a beam of light strikes an object, the beam could be reflected, absorbed or transmitted. The beam of light when irradiated on the object reflects back one portion of light from the outside surface of the object called the specular reflectance and this portion of light does not acquire any information from the object. One other portion of the light penetrates the object and is then reflected back from within the object. This is called diffuse reflectance where the light is reflected in many directions giving the object a matt finish. For example, specular reflectance is the process that prevents the observer from seeing through a window on a sunny day and diffuse reflectance is the process that allows the observer to view into the house when the eyes are shaded with hands. Transmittance is the ratio of light that is transmitted through the object to that portion of light that impinges on the object. In the field of NIR spectroscopy, absorbance is the logarithm of the reciprocal of transmittance (Williams and Norris 1987).

The spectrophotometer is an instrument that measures flux of light at specified wavelengths and reflectance characteristic is the ratio of the flux that is reflected from the object to the flux that a standard reflective surface would reflect. The standard surface used is a reference disc made of polytetrafluroethylene, which exhibits NIR performance while maintaining UV and Visible performance. The spectrophotometer has a diffuse reflectance accessory (DRA), which consists of a 110 mm diameter integrating sphere with the ability of collecting most of the reflected light removing any directional preferences and presenting an integrated signal to the detector. The schematic diagram of the DRA is illustrated in Fig 2.1. The sphere in turn consists of an in-built high performance photomultiplier tube to pick up signals in the ultraviolet and the visible regions and a lead sulphide detector to collect signals in the near infrared region of the electromagnetic spectrum. There is a set of mirrors, which facilitates the focusing of the light beam on to the sample, which is usually placed in a powder cell (Fig. 2.2). The spectrophotometer has two bulbs, one a visible IQ bulb emitting light ranging from the infrared to the visible region and the other a deuterium bulb emitting light in the UV region. These bulbs are mounted on a rotating plate placed inside a container to help block the stray light. The plate positions itself according to the wavelength requirement.

When actuated, a beam of light passes from the bulb through a 10 mm opening in the container and through a lens where the beam is focused. It then strikes a mirror, which has gratings equaling 1200 lines/mm on one side for the infrared region and 700 lines/mm for the visible and UV regions to facilitate obtaining the required wavelength. It then passes through preset filters to fine tune the beam to the wavelength requirement. Following this, the beam strikes another set of mirrors and is split into two, one being the reference beam and the other the sample beam. The reference beam is calibrated against a base line and a zero line collected after setting the scan parameters using the standard reference disc. The reference beam is diffused directly into the sphere via a reference port before being measured by the detector. The sample beam on the other hand passes on to the set of mirrors (M_1, M_2) and then to an offset lens (Fig. 2.1). Once it passes through the offset lens, it is focused on to the sample port where the sample is placed. The beam is then diffused through the integrating sphere. The signal from the reference beam and the signal from the sample beam are integrated as one and presented to the appropriate detectors.

The manufacturer claims the accuracy of the spectrophotometer could be affected mainly because of the deterioration of the reference disc or the coating of the integrating sphere from contact with dust, fingerprints and smoke, aging of light, incorrectly placing the samples at the port and letting the diffused light to escape.

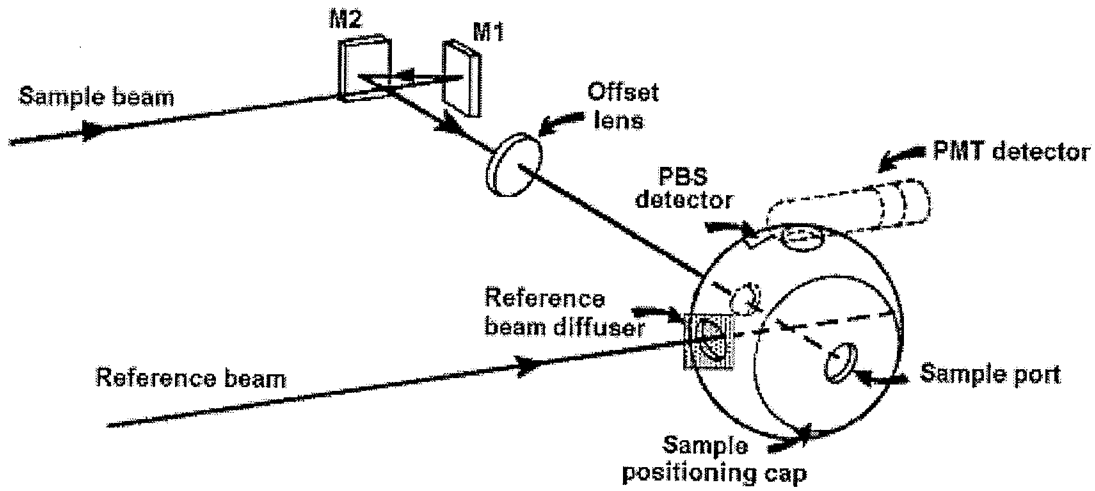

Fig 2.1 Schematic diagram of the diffuse reflectance accessory
(Anonymous 2001)

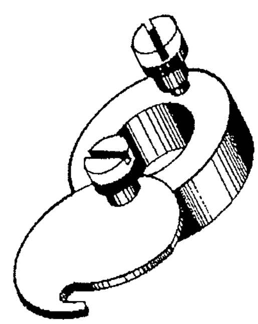

Fig 2.2 Schematic diagram of the powder cell
(Anonymous 2001)

3: MATERIALS AND METHODS

3.1 Spectrophotometer

The equipment consisted of a spectrophotometer (Model: Cary 5, Varian Canada Inc., Mississauga, ON) with a diffuse reflectance accessory comprising an integrating sphere with five ports and coated with polytetrafluroethylene. It also had two detectors: a photomultiplier tube that collected signals in the visible and the ultraviolet regions and a lead-sulphide detector to collect signals in the near infrared regions. The change in detectors plus a change in the side of the mirror with gratings introduced a shift in the reflectance recordings at 800 nm, which was counteracted by fixing the change in detectors at 870 nm. A personal computer loaded with "Varian", a windows based software, acted as an interface with the spectrophotometer. The machine was calibrated using the standard reference material and recording a base line and a zero reflectance line. Total reflectance, which comprises both diffuse and specular reflectance, was also recorded.

3.2 Bulk Grain Samples

Bulk grain samples of seven types of cereals and buckwheat, three types of oilseeds, ten types of pulses, and 26 types of specialty seeds were procured. Buckwheat was lumped with cereals because it is a common contaminant of cereals. The cereal grains were: Canada Western Red Spring wheat, Canada Western Amber Durum wheat, Soft White Spring Wheat, 2 – row barley, 6 – row barley, oats, and rye. The oilseeds were: sunflower, canola, and flaxseeds. The pulse seeds were: small red kidney beans, light red kidney beans, black beans, navy pea beans, pinto beans, Eston lentils, Laird lentils, dark speckled green lentils, Espace green peas, and Croma yellow peas. The specialty seeds were: alfalfa, alsike clover, annual rye grass, birds foot trefoil, brown mustard, creeping bent grass, creeping red fescue, crested wheat grass, crown millet, intermediate wheat grass, Kentucky blue grass, meadow brome grass, meadow fescue, orchard grass, oriental mustard, perennial rye grass, red clover, reed canary grass, Siberian millet, slender wheat grass, sorghum Sudan grass, sweet clover, tall fescue, timothy, and yellow mustard. The samples were kept in a freezer at -18° C and were equilibrated to room temperature before use.

The CWRS wheat was collected, from 15 different growing regions in western Canada from various years, to represent different growing conditions and crop-years to test the effect of, the different growing environments and crop-age on the variables measured. The CWRS wheat samples were collected from sub-boreal (four), sub-humid-prairie (eight), and semi arid regions (three) based on the climatic subdivisions of the Canadian prairies (Paliwal 2002; Putnam and Putnam 1970). The CWRS wheat samples consisted of 11 samples from 1998, two from 1999, and two from 2000 (Fig. 3.1, 3.2, and 3.3).

The wheat sample from North Battleford, SK was conditioned to 10, 12, 14, 16, and 18 % nominal moisture contents. The samples were conditioned to the required moisture content by adding the appropriate amount of water, mixing it thoroughly and leaving the sample airtight for a period of 12 h for the moisture to penetrate uniformly

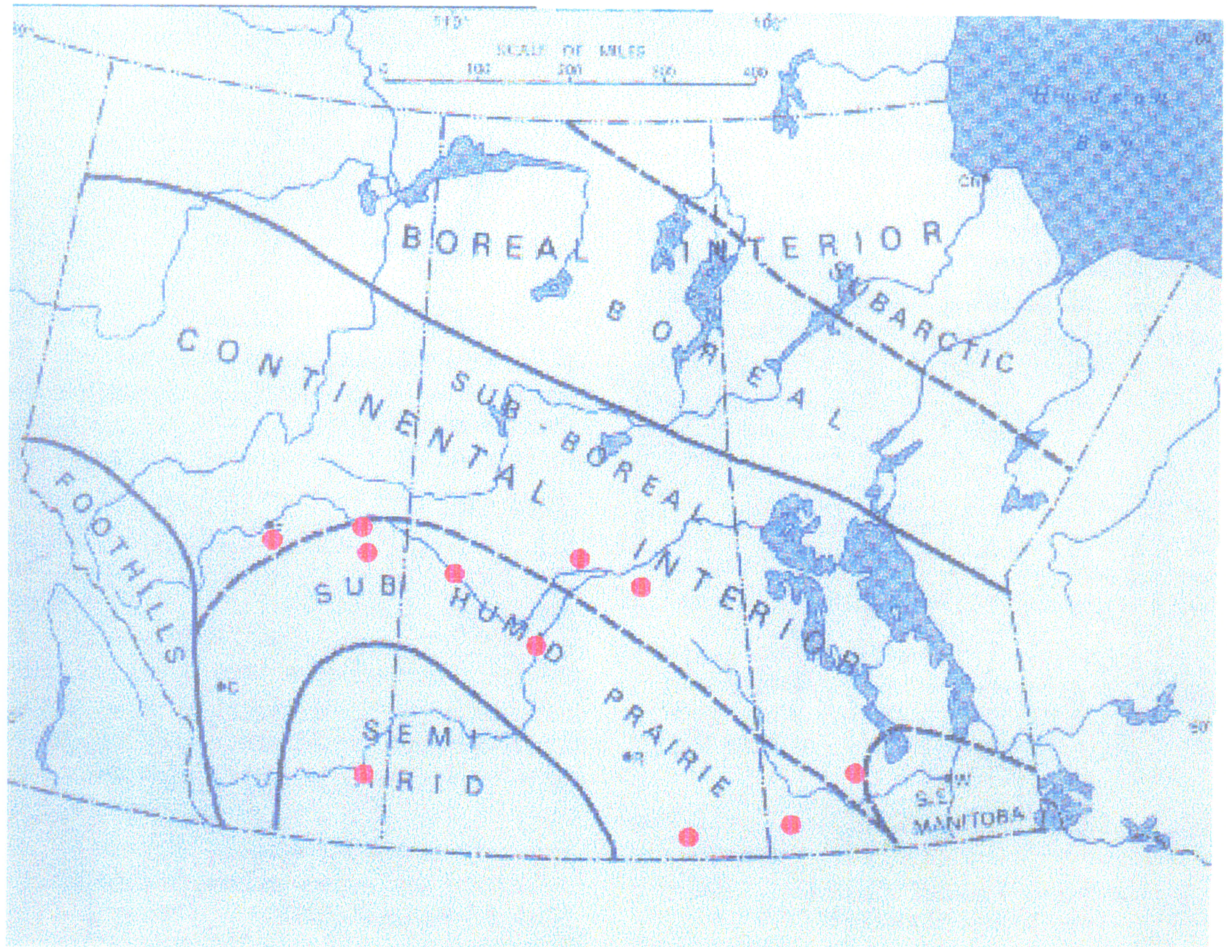

Fig. 3.1 Sample locations for the CWRS wheat samples collected in 1998, representing the different climatic conditions (Putnam and Putnam 1970)

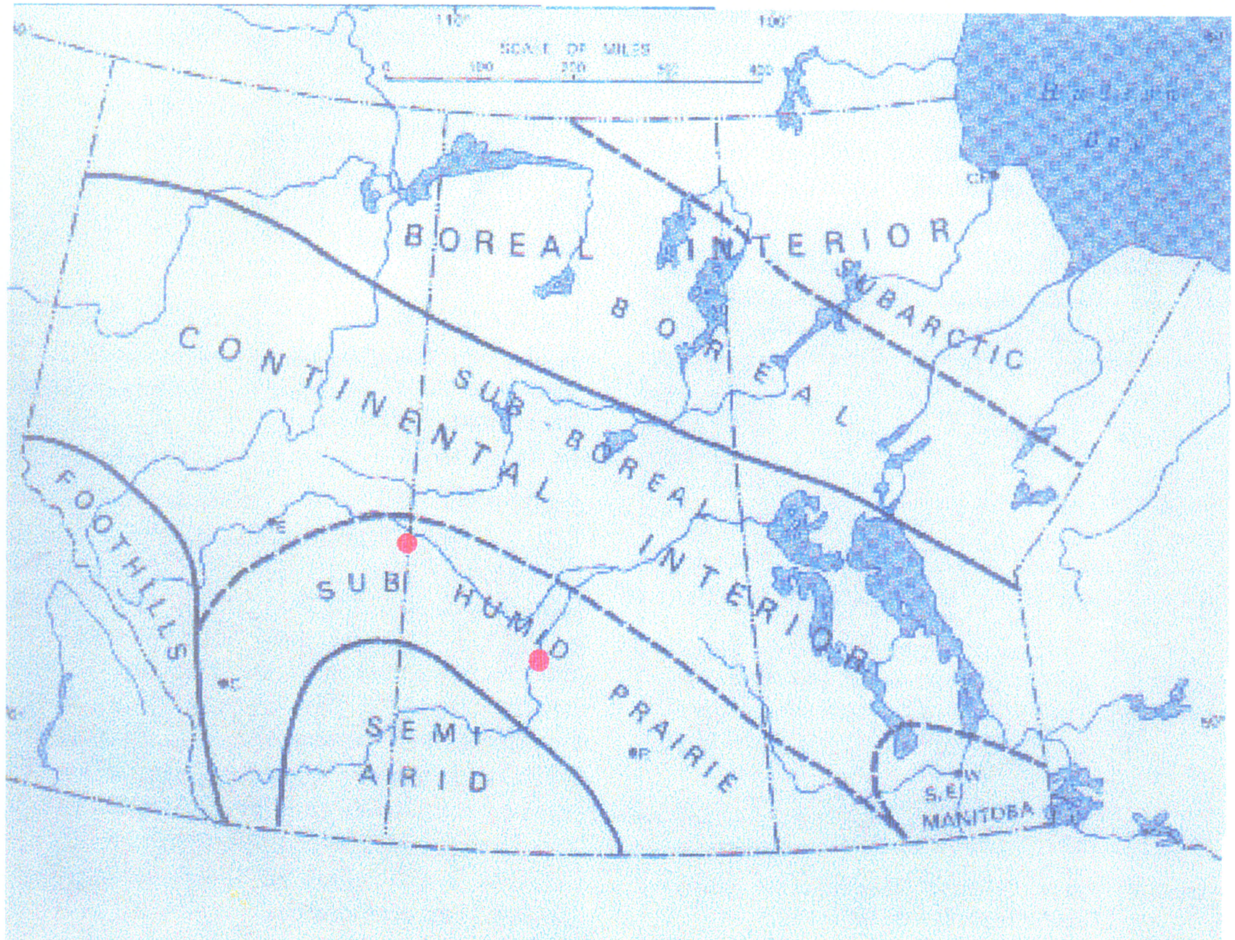

Fig 3.2 Sample locations for the CWRS wheat samples collected in 1999 (Putnam and Putnam)

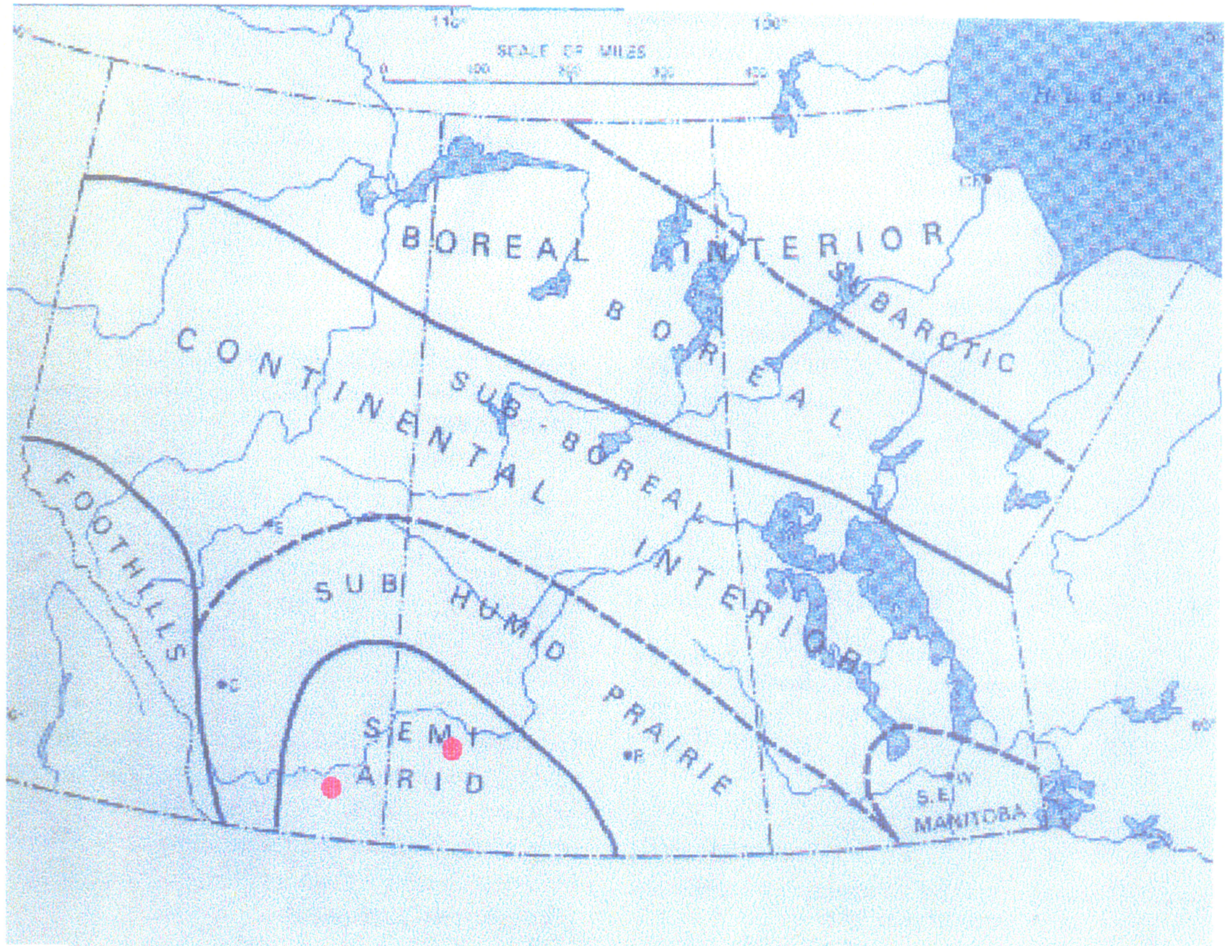

Fig. 3.3 Sample locations for the CWRS wheat samples collected in 2000 (Putnam and Putnam)

through the grain sample. The final moisture contents of the prepared samples were 9.8, 11.8, 14.1, 15.9 and 17.7 %, respectively. Moisture contents of the wheat samples were determined by drying triplicate samples of whole seeds (10 g each) at 130° C for 19 h and obtaining the difference in weight of the samples before and after drying (ASAE 1997).

The wheat sample from Estevan, SK was conditioned to have 3, 6, 9, 12, and 15 % foreign material. Appropriate amounts of foreign material to the hand-cleaned wheat samples to produce the desired mixture. (e.g., 3 g of foreign material was added to 97 g of pure wheat to give a 3 % mixture) .

3.3 Sampling and Analysis

Samples of about 500 g of all the tested grains were obtained. A sub-sample weighing 3-5 g was randomly selected, packed tightly into the powder cell and fixed to the sample port of the integrating sphere in the diffuse reflectance accessory. The spectrophotometer collected the percent reflectance data at a preset interval of 0.33 nm over a range of 320 nm to 1880 nm three times. The values were averaged and plotted as one point for each nanometer of the wavelength. The data were filtered using a preset filter to account for the disturbance and noise and plotted over an interval of 5 nm for the entire range. The variation in an undisturbed sample was within ± 0.3 %. Allowing two volunteers to test the setting of the powder cell in the sample port tested user-to-user variation and no significant difference was found as the integrated sphere was marked for the positioning of the sample holder. The procedure of randomly selecting and testing the samples was repeated 30 times. Each time the sample was picked from a container, tested and placed back into the container before the procedure was repeated. A total of 81 samples from different seed types were tested, each replicated 30 times and the data collected were used for the analysis. A typical variation within 30 replicates of the same sample of CWRS wheat collected from Churchill, MB is shown in Fig. 3.4.

The percent reflectance data thus collected cannot be used in real-time industrial applications because the data may be affected by light intensity, dust, image background and aging of light source (Majumdar et al. 1996). A more effective approach to utilize these data would be to use the percent reflectance ratios, slopes, and the ratio of slopes. Since there were variations in the reflectance data along the whole wavelength range for different grains (in contrast to Eu (1997) and Hawk et al. (1970) who noted differences in certain regions) it was decided to calculate the three parameters using the data along the entire range with a 10 nm interval.

A total of 465 features were extracted from the percent reflectance data, 156 features for the ratio model, 155 for the slope model and 154 for the slope-ratio model were used for analysis. The features were defined as follows:

Feature 1 : Percent reflectance ratio = % Reflectance at 320nm / % Reflectance at 330nm

Feature 2 : Percent reflectance ratio = % Reflectance at 330nm / % Reflectance at 340nm

.
.
.
.

Feature 155: Percent reflectance ratio = % Reflectance at 1860nm / % Reflectance at 1870nm

Feature 156: Percent reflectance ratio = % Reflectance at 1870nm / % Reflectance at 1880 nm

Feature 157: Percent reflectance slope = $\frac{\text{\%Reflectance at 330nm- \% Reflectance at 320nm}}{10}$

Feature 158: Percent reflectance slope = $\frac{\text{\%Reflectance at 340nm-\% Reflectance at 330nm}}{10}$

$\vdots$

Feature 310: Percent reflectance slope = $\frac{\text{\%Reflectance at 1870nm-\%Reflectance at 1860nm}}{10}$

Feature 311: Percent reflectance slope = $\frac{\text{\% Reflectance at 1880nm-\% Reflectance at 1870nm}}{10}$

Feature 312 : Ratio of slope = $\frac{\text{\% Reflectance at 330nm - \% Reflectance at 320nm}}{\text{\% Reflectance at 340nm - \% Reflectance at 330nm}}$

Feature 313 : Ratio of slope = $\frac{\text{\% Reflectance at 340nm - \% Reflectance at 330nm}}{\text{\% Reflectance at 350nm - \% Reflectance at 340nm}}$

$\vdots$

Feature 464 : Ratio of slope = $\frac{\text{\% Reflectance at 1860nm- \% Reflectance at 1850nm}}{\text{\% Reflectance at 1870nm- \% Reflectance at 1860nm}}$

Feature 465 : Ratio of slope = $\frac{\text{\% Reflectance at 1870nm- \% Reflectance at 1860nm}}{\text{\% Reflectance at 1880nm- \% Reflectance at 1870nm}}$

Procedure STEPDISC (SAS 1990) was used to test the contribution of each of the features in the three models. The best 20 features were selected and were used in assessing the classification accuracies of the models using Procedure DISCRIM. The twenty-four replicates out of the thirty replicates were selected randomly and used for the training set and the remaining six replicates were used as the test set. This procedure was repeated thrice for three different training and test sets.

Back Propagation Neural Network (BPNN) was another classifier used to test the classification by reflectance data using the three models. Jayas et al. (2000) reported that the most popular choice for the classification of agricultural products would be the BPNN. The neural network was implemented using a software package called NeuroShell 2 developed by Ward Systems Group, Fredrick, MD. It consists of one input layer, two hidden layers and one output layer. The network is trained until it reaches a predefined number of learning epochs (Paliwal 2002, Paliwal et al. 2001Visen 2002). Here the network displays a weight for each of the input variables, which is a measure of the contribution of each input variable to the classification. The top 20 features were used and tested against the same three test sets that were used for testing the classification accuracy of the discriminant analysis.

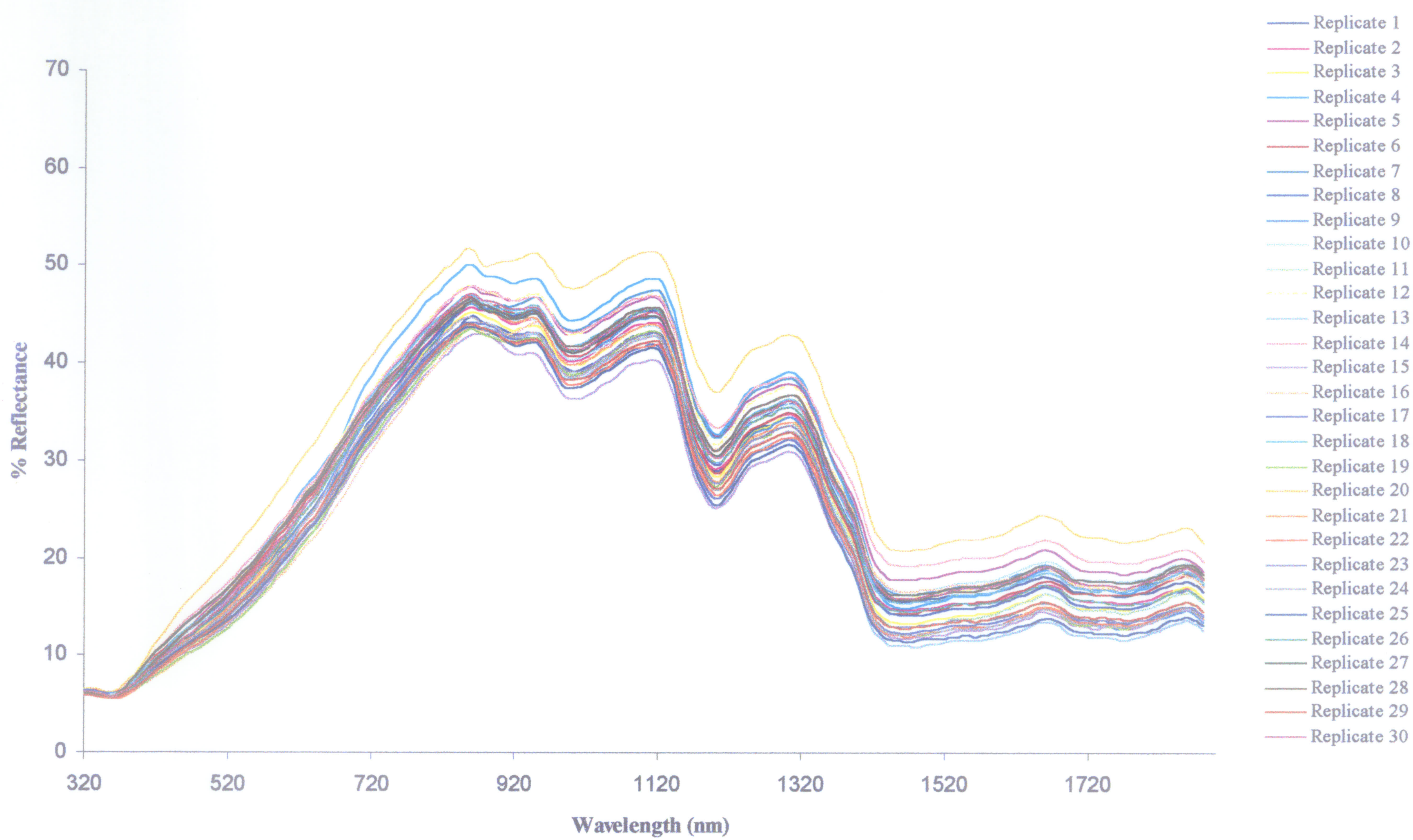

Fig. 3.4 Reflectance characteristics of CWRS wheat showing the variability in measurements based on 30 replicates of the same sample

4. RESULTS AND DISCUSSION

4.1 Cereals and Buckwheat

From the reflectance data (Fig. A1, Appendix A) of cereals, the ratio, slope and slope ratio features were extracted. The contributions of the individual features to the classifiers were determined and the features were ranked in descending order (Tables 4.1, 4.2, and 4.3) for the ratio, slope and slope ratio models, respectively. Using the top 20 features for each model, classifications were performed for the cereals and the results are given in Tables 4.4, 4.5, and 4.6, respectively.

The features listed in the tables show that the top 20 features are not limited to one individual region but are from the whole scanned range. The top 20 features in both the classifiers in each of the three models (Tables 4.1, 4.2, and 4.3) are not the same. The differences are due to the different working procedure of the two classifiers (Paliwal 2002).

A classification accuracy of 100 % was obtained for the seven cereals and buckwheat using the ratio and the slope ratio models (Tables 4.4 and 4.6, respectively). A classification accuracy of 100 % was obtained for six of the seven cereals and buckwheat using the slope model (Table 4.5). CWRS wheat was once misclassified as CWAD wheat resulting in 94.4 % accuracy when non-parametric estimation was used and SWSW wheat was once misclassified as CWAD wheat resulting in 94.4 % accuracy when BPNN was used (Table 4.5).

Though the results show that the reflectance characteristics are highly successful in classifying the cereals, the use of the top twenty features by each of the classifiers where the features lie over the entire range (320 – 1880 nm) makes it impossible to use reflectance characteristics in real time applications. To have a spectrophotometer at a primary or terminal elevator would be expensive. It was hoped that a few selected wavelengths would be able to classify cereals and buckwheat and filters to acquire images at these wavelengths could be used. Other procedures such as digital image analysis of bulk samples may be a better choice for industrial applications (Majumdar and Jayas 1999, Paliwal 2002, Visen 2002).

Table 4.1 Top 20 features of the ratio model using STEPDISC and BPNN for cereals and buckwheat

Rank	Selected features for hold-out method	Average squared canonical correlation	Partial r^2	Selected features for BPNN	Weight
1	Feature 45	0.139	0.974	Feature 10	0.009
2	Feature 90	0.274	0.955	Feature 9	0.008
3	Feature 10	0.398	0.915	Feature 96	0.008
4	Feature 111	0.527	0.906	Feature 8	0.008
5	Feature 16	0.638	0.844	Feature 89	0.007
6	Feature19	0.756	0.842	Feature 30	0.007
7	Feature 108	0.789	0.722	Feature 5	0.007
8	Feature 5	0.862	0.647	Feature 20	0.007
9	Feature 86	0.869	0.480	Feature 136	0.007
10	Feature 100	0.886	0.404	Feature 122	0.007
11	Feature 102	0.894	0.320	Feature 6	0.007
12	Feature 81	0.903	0.347	Feature 137	0.007
13	Feature 21	0.909	0.271	Feature 16	0.007
14	Feature 33	0.914	0.271	Feature 121	0.007
15	Feature 56	0.919	0.243	Feature 69	0.007
16	Feature 137	0.921	0.236	Feature 155	0.007
17	Feature 98	0.924	0.230	Feature 142	0.007
18	Feature 151	0.928	0.208	Feature 29	0.007
19	Feature 144	0.932	0.250	Feature 26	0.007
20	Feature 87	0.934	0.189	Feature 102	0.007

Table 4.2 Top 20 features of the slope model using STEPDISC and BPNN for cereals and buckwheat

Rank	Selected features for hold-out method	Average squared canonical correlation	Partial r^2	Selected features for BPNN	Weight
1	Feature 222	0.138	0.968	Feature 166	0.010
2	Feature 169	0.274	0.957	Feature 278	0.009
3	Feature 267	0.405	0.920	Feature 172	0.008
4	Feature 185	0.525	0.879	Feature 165	0.008
5	Feature 175	0.625	0.779	Feature 256	0.008
6	Feature 173	0.726	0.792	Feature 168	0.008
7	Feature 162	0.817	0.721	Feature 300	0.008
8	Feature 293	0.838	0.675	Feature 289	0.008
9	Feature 177	0.855	0.426	Feature 189	0.007
10	Feature 245	0.876	0.411	Feature 167	0.007
11	Feature 312	0.881	0.374	Feature 293	0.007
12	Feature 201	0.891	0.372	Feature 250	0.007
13	Feature 166	0.900	0.361	Feature 260	0.007
14	Feature 213	0.903	0.343	Feature 285	0.007
15	Feature 192	0.909	0.327	Feature 311	0.007
16	Feature 253	0.911	0.283	Feature 201	0.007
17	Feature 212	0.918	0.283	Feature 169	0.007
18	Feature 289	0.921	0.260	Feature 164	0.007
19	Feature 255	0.923	0.246	Feature 276	0.007
20	Feature 165	0.925	0.200	Feature 291	0.010

Table 4.3 Top 20 features of the slope-ratio model using STEPDISC and BPNN for cereals and buckwheat

Rank	Selected features for hold-out method	Average squared canonical correlation	Partial r^2	Selected features for BPNN	Weight
1	Feature 403	0.049	0.990	Feature 331	0.008
2	Feature 349	0.098	0.981	Feature 330	0.008
3	Feature 398	0.146	0.972	Feature 401	0.008
4	Feature 325	0.195	0.967	Feature 447	0.008
5	Feature 347	0.242	0.967	Feature 349	0.007
6	Feature 352	0.290	0.952	Feature 323	0.007
7	Feature 314	0.337	0.941	Feature 375	0.007
8	Feature 395	0.382	0.924	Feature 412	0.007
9	Feature 336	0.426	0.904	Feature 326	0.007
10	Feature 423	0.467	0.888	Feature 397	0.007
11	Feature 341	0.510	0.892	Feature 352	0.007
12	Feature 401	0.546	0.859	Feature 459	0.007
13	Feature 356	0.587	0.854	Feature 328	0.007
14	Feature 419	0.613	0.781	Feature 319	0.007
15	Feature 414	0.645	0.781	Feature 337	0.007
16	Feature 323	0.681	0.760	Feature 340	0.007
17	Feature 318	0.708	0.722	Feature 353	0.007
18	Feature 381	0.735	0.652	Feature 385	0.007
19	Feature 331	0.753	0.575	Feature 411	0.007
20	Feature 377	0.762	0.555	Feature 413	0.007

Table 4.4 Classification accuracies of the cereals and buckwheat using the ratio model in non- parametric estimation and BPNN

Cereal	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
2 row barley				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
6 row barley				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
CWRS wheat				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Buck wheat				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
CWAD wheat				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Oats				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Rye				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
SWSW				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

Table 4.5 Classification accuracies of the cereals and buckwheat using the slope model in non- parametric estimation and BPNN

Cereal	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
2 row barley				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
6 row barley				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
CWRS wheat				
Set 1	6		6	
Set 2**	5	94.4± 9.6	6	100
Set 3	6		6	
Buck wheat				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
CWAD wheat				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Oats				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Rye				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
SWSW				
Set 1	6		6	
Set 2***	6	100	5	94.4± 9.6
Set 3	6		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

** CWRS wheat was misclassified as CWAD wheat.

*** SWSW was misclassified as CWAD wheat.

Table 4.6 Classification accuracies of the cereals and buckwheat using the slope-ratio model in non- parametric estimation and BPNN

Cereal	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
2 row barley				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
6 row barley				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
CWRS wheat				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Buck wheat				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
CWAD wheat				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Oats				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Rye				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
SWSW				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

4.2 Oilseeds

From the reflectance data (FigA2, Appendix A) ratio, slope, and slope-ratio features were extracted. The contributions of the individual features to the classifiers were determined and the features were ranked in descending order (Tables 4.7, 4.8, and 4.9), respectively, for the three models. Using the top 20 features for each model, classifications were done for oilseeds and results are given in Tables 4.10, 4.11, and 4.12, respectively for the ratio, slope and, slope ratio models. The oilseeds were correctly (100%) classified using the three models for both the classifiers (Tables 4.10, 4.11, and 4.12).

The reflectance characteristics in Fig. A2 show the wide variation between the three oilseeds and is an indicator of the success in the classification process. It can be seen that the classifiers are not using the same features for the classification process, as explained previously because the methods adopted by them for classification are different. The selected features are scattered throughout the scanned range making it impossible to use filters to acquire reflectance at selected wavelengths. Table 4.9 shows the high partial r^2 value in the slope ratio model compared to the other two models (Tables 4.7, and 4.8). This means that the slope-ratio model is a more robust model compared to the other two models.

Table 4.7 Top 20 features of the ratio model using STEPDISC and BPNN for oilseeds

Rank	Selected features for hold-out method	Average squared canonical correlation	Partial r^2	Selected features for BPNN	Weight
1	Feature 98	0.226	0.905	Feature 143	0.008
2	Feature 100	0.262	0.549	Feature 109	0.007
3	Feature 53	0.346	0.409	Feature 66	0.007
4	Feature 97	0.370	0.346	Feature 133	0.007
5	Feature 89	0.430	0.306	Feature 118	0.007
6	Feature 87	0.441	0.301	Feature 9	0.007
7	Feature 1	0.466	0.297	Feature 139	0.007
8	Feature 144	0.485	0.251	Feature 22	0.007
9	Feature 69	0.521	0.186	Feature 28	0.007
10	Feature 62	0.551	0.192	Feature 145	0.007
11	Feature 68	0.583	0.166	Feature 83	0.007
12	Feature 155	0.595	0.166	Feature 23	0.007
13	Feature 99	0.602	0.156	Feature 45	0.007
14	Feature 103	0.622	0.149	Feature 89	0.007
15	Feature 90	0.630	0.143	Feature 123	0.007
16	Feature 121	0.646	0.136	Feature 101	0.007
17	Feature 71	0.658	0.123	Feature 125	0.007
18	Feature 92	0.671	0.140	Feature 136	0.007
19	Feature 64	0.678	0.131	Feature 149	0.007
20	Feature 69	0.685	0.125	Feature 87	0.007

Table 4.8 Top 20 features of the slope model using STEPDISC and BPNN for oilseeds

Rank	Selected features for hold-out method	Average squared canonical correlation	Partial r^2	Selected features for BPNN	Weight
1	Feature 239	0.497	0.994	Feature 299	0.008
2	Feature 244	0.988	0.984	Feature 265	0.007
3	Feature 248	0.993	0.550	Feature 222	0.007
4	Feature 190	0.995	0.454	Feature 274	0.007
5	Feature 245	0.995	0.326	Feature 296	0.007
6	Feature 177	0.996	0.199	Feature 184	0.007
7	Feature 299	0.996	0.168	Feature 296	0.007
8	Feature 255	0.996	0.101	Feature 174	0.007
9	Feature 293	0.997	0.103	Feature 186	0.007
10	Feature 291	0.997	0.091	Feature 179	0.007
11	Feature 243	0.997	0.072	Feature 301	0.007
12	Feature 157	0. 997	0.066	Feature 279	0.007
13	Feature 212	0. 997	0.075	Feature 251	0.007
14	Feature 185	0. 997	0.072	Feature 245	0.007
15	Feature 201	0. 997	0.080	Feature 279	0.007
16	Feature 205	0. 997	0.090	Feature 257	0.007
17	Feature 192	0. 997	0.063	Feature 281	0.007
18	Feature 188	0. 997	0.140	Feature 292	0.007
19	Feature 197	0.998	0.143	Feature 305	0.007
20	Feature 173	0.998	0.061	Feature 243	0.007

Table 4.9 Top 20 features of the slope-ratio model using STEPDISC and BPNN for oil seeds

Rank	Selected features for hold-out method	Average squared canonical correlation	Partial r^2	Selected features for BPNN	Weight
1	Feature 468	0.023	1.000	Feature 421	0.008
2	Feature 412	0.046	1.000	Feature 334	0.008
3	Feature 390	0.069	1.000	Feature 445	0.007
4	Feature 465	0.093	1.000	Feature 455	0.007
5	Feature 321	0.111	1.000	Feature 430	0.007
6	Feature 426	0.134	1.000	Feature 437	0.007
7	Feature 367	0.158	1.000	Feature 463	0.007
8	Feature 378	0.181	1.000	Feature 457	0.007
9	Feature 407	0.204	1.000	Feature 340	0.007
10	Feature 452	0.227	1.000	Feature 378	0.007
11	Feature 376	0.250	1.000	Feature 335	0.007
12	Feature 458	0. 274	1.000	Feature 404	0.007
13	Feature 436	0. 296	1.000	Feature 435	0.007
14	Feature 377	0. 319	1.000	Feature 401	0.007
15	Feature 320	0. 342	1.000	Feature 395	0.007
16	Feature 326	0. 365	1.000	Feature 389	0.007
17	Feature 366	0. 388	1.000	Feature 347	0.007
18	Feature 427	0. 411	1.000	Feature 400	0.007
19	Feature 440	0.433	1.000	Feature 452	0.007
20	Feature 319	0.455	1.000	Feature 332	0.007

Table 4.10 Classification accuracies of the oilseeds using the ratio model in non- parametric estimation and BPNN

Oilseeds	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
Canola				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Flax seed				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Sunflower				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

Table 4.11 Classification accuracies of the oilseeds using the slope model in non- parametric estimation and BPNN

Oilseeds	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
Canola				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Flax seed				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Sunflower				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

Table 4.12 Classification accuracies of the oilseeds using the slope-ratio model in non- parametric estimation and BPNN

Oilseeds	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
Canola				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Flax seed				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Sunflower				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

4.3 Pulses

From the reflectance data (FigA3a, and A3b, Appendix A) ratio, slope and slope-ratio features were extracted. The contribution of the individual features to the classifiers was determined and the features were ranked in descending order (Tables 4.13, 4.14, and 4.15), respectively, for the three models. Using the top 20 features for each model, classifications were done for pulses and results are given in Tables 4.16, 4.17 and, 4.18, respectively for the ratio, slope, and slope ratio models.

The pulses were correctly (100%) classified using the three models for both the classifiers (Tables 4.16, 4.17, and 4.18). The reflectance data (Fig. A3a and A3b, Appendix A) show a variation in the visible region among all the pulses processed and this is reflected in the feature selection by the classifiers for the classification process (Tables 4.13, 4.14, and 4.15).

The statistical values of the features selected prove, that the slope-ratio feature is a more effective model than the other two models although the classification was possible with all the three models (Tables 4.16, 4.17, and 4.18).

Although the features selected by the classifiers are different from one another in the three models (ratio, slope, and slope ratio models) the concentration of features is in the visible spectrum with the exception of a few features in the near infrared region closer to the visible region (Tables 4.13, 4.14, and 4.15).

As explained in the previous case, the difficulty of incorporating these models in an industrial setting makes it impractical and limits its use to laboratory conditions.

Table 4.13 Top 20 features of the ratio model using STEPDISC and BPNN for pulses

Rank	Selected features for hold-out method	Average squared canonical correlation	Partial r^2	Selected features for BPNN	Weight
1	Feature 32	0.226	0.982	Feature 30	0.009
2	Feature 48	0.262	0.968	Feature 36	0.009
3	Feature 40	0.346	0.948	Feature 41	0.008
4	Feature 37	0.370	0.945	Feature 34	0.008
5	Feature 28	0.430	0.816	Feature 29	0.008
6	Feature 98	0.441	0.753	Feature 32	0.008
7	Feature 1	0.466	0.720	Feature 100	0.008
8	Feature 5	0.485	0.623	Feature 9	0.008
9	Feature 50	0.521	0.624	Feature 40	0.008
10	Feature 115	0.551	0.507	Feature 42	0.008
11	Feature 11	0.583	0.503	Feature 21	0.008
12	Feature 39	0.595	0.455	Feature 28	0.008
13	Feature 9	0.602	0.455	Feature 76	0.008
14	Feature 36	0.622	0.426	Feature 133	0.008
15	Feature 112	0.630	0.477	Feature 137	0.007
16	Feature 156	0.646	0.433	Feature 8	0.007
17	Feature 100	0.658	0.542	Feature 33	0.007
18	Feature 25	0.671	0.347	Feature 2	0.007
19	Feature 66	0.678	0.333	Feature 5	0.007
20	Feature 35	0.685	0.305	Feature 20	0.007

Table 4.14 Top 20 features of the slope model using STEPDISC and BPNN for pulses

Rank	Selected features for hold-out method	Average squared Canonical correlation	Partial r^2	Selected features for BPNN	Weight
1	Feature 190	0.109	0.982	Feature 190	0.009
2	Feature 196	0.214	0.959	Feature 198	0.009
3	Feature 207	0.320	0.953	Feature 189	0.008
4	Feature 208	0.423	0.948	Feature 197	0.008
5	Feature 157	0.521	0.908	Feature 231	0.008
6	Feature 202	0.593	0.830	Feature 192	0.008
7	Feature 182	0.680	0.804	Feature 256	0.008
8	Feature 243	0.749	0.786	Feature 232	0.008
9	Feature 194	0.790	0.719	Feature 236	0.008
10	Feature 175	0.837	0.577	Feature 293	0.008
11	Feature 231	0.847	0.513	Feature 185	0.008
12	Feature 254	0.859	0.490	Feature 186	0.008
13	Feature 256	0.871	0.377	Feature 161	0.007
14	Feature 232	0.879	0.360	Feature 311	0.007
15	Feature 271	0.888	0.351	Feature 172	0.007
16	Feature 170	0.898	0.339	Feature 165	0.007
17	Feature 166	0.901	0.349	Feature 188	0.007
18	Feature 263	0.907	0.377	Feature 310	0.007
19	Feature 267	0.913	0.307	Feature 177	0.007
20	Feature 193	0.916	0.254	Feature 254	0.007

Table 4.15 Top 20 features of the slope-ratio model using STEPDISC and BPNN for pulses

Rank	Selected features for hold-out method	Average squared canonical correlation	Partial r^2	Selected features for BPNN	Weight
1	Feature 409	0.013	0.996	Feature 350	0.011
2	Feature 333	0.026	0.994	Feature 348	0.011
3	Feature 338	0.038	0.982	Feature 355	0.011
4	Feature 444	0.051	0.976	Feature 349	0.010
5	Feature 326	0.063	0.958	Feature 375	0.010
6	Feature 373	0.075	0.960	Feature 320	0.010
7	Feature 459	0.087	0.925	Feature 354	0.010
8	Feature 447	0.098	0.904	Feature 344	0.009
9	Feature 362	0.109	0.884	Feature 330	0.009
10	Feature 342	0.120	0.881	Feature 347	0.009
11	Feature 318	0.131	0.864	Feature 351	0.009
12	Feature 464	0.141	0.846	Feature 315	0.009
13	Feature 344	0.149	0.845	Feature 332	0.009
14	Feature 339	0.158	0.843	Feature 396	0.009
15	Feature 325	0.166	0.839	Feature 357	0.009
16	Feature 378	0.175	0.794	Feature 331	0.008
17	Feature 336	0.183	0.800	Feature 346	0.008
18	Feature 375	0.193	0.897	Feature 352	0.008
19	Feature 426	0.200	0.816	Feature 377	0.008
20	Feature 445	0.208	0.821	Feature 322	0.008

Table 4.16 Classification accuracies of the pulses using the ratio model in non- parametric estimation and BPNN

Pulses	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean(%)*
Black beans				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Croma yellow peas				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Dark speckled green lentils				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Espace Green peas				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Eston lentils				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Laird lentils				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Light red kidney beans				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Navy beans				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Pinto beans				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Small reds				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

Table 4.17 Classification accuracies of the pulses using the slope model in non- parametric estimation and BPNN

Pulses	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
Black beans				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Croma yellow peas				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Dark speckled green lentils				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Espace Green peas				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Eston lentils				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Laird lentils				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Light red kidney				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Navy beans				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Pinto beans				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Small reds				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

Table 4.18 Classification accuracies of the pulses using the slope-ratio model in non- parametric estimation and BPNN

Pulses	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
Black beans				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Croma yellow peas				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Dark speckled green lentils				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Espace Green peas				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Eston lentils				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Laird lentils				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Light red kidney beans				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Navy beans				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Pinto beans				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Small reds				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

4.4 Cereals and Buckwheat, Oilseeds, and Pulses

From the reflectance data (FigA1, A2, A3a, and A3b, Appendix A) ratio, slope and slope-ratio features were extracted. The contributions of the individual features to the classifiers were determined and the features were ranked in descending order (Tables 4.19, 4.20, and 4.21). Using the top 20 features for each model, classifications were done for the combination of cereals, oilseeds and pulses and results are given in Tables 4.22, 4.23, and 4.24, respectively, for the ratio, slope, and slope ratio models.

The tested cereals, oilseeds, and pulses were correctly (100%) classified using the slope ratio model for both the classifiers (Tables 4.24). The ratio model, however, misclassified three of the pulses namely Croma yellow peas, dark green speckled lentils, and pinto beans when the statistical classifier was used but correctly (100%) classified using the BPNN classifier. The slope model misclassified one cereal and three pulses using the statistical classifier and one cereal and five pulses when the BPNN classifier was used. The confusion matrices for classification accuracies are presented in Tables B1a, B1b, B1c, B1d, B1e, and B1f, Appendix B. The slope-ratio model, similar to the previous sections, was more successful in classifying the samples than the ratio or the slope model.

The ratio model was, however, better compared to the slope model. The classification accuracies in the ratio and slope model could be greatly improved by initially classifying each sample to its respective group, i.e. cereals, oilseeds or pulses, and then using the features for that group given in the tables in the previous three sections to individual classes and achieving a 100% accuracy in the ratio and the slope model. The confusion matrices (Tables B1a, and B1b) show very clearly how the misclassification is spread into the specific groups and helps in understanding the nature of the problem.

The slope-ratio model, which had the best classification accuracy, uses the features mostly from the near-infra red region unlike the slope model, which uses features from the visible, and UV region (Tables 4.19, 4.20, and 4.21).

Table 4.19 Top 20 features of the ratio model using STEPDISC and BPNN for Cereals and buckwheat, oilseeds and pulses

Rank	Selected features for hold-out method	Average squared canonical correlation	Partial r^2	Selected features for BPNN	Weight
1	Feature 138	0.050	0.992	Feature 155	0.010
2	Feature 50	0.099	0.980	Feature 8	0.009
3	Feature 86	0.147	0.977	Feature 9	0.009
4	Feature 40	0.195	0.958	Feature 29	0.009
5	Feature 82	0.243	0.957	Feature 7	0.008
6	Feature 37	0.290	0.950	Feature 41	0.008
7	Feature 112	0.336	0.938	Feature 98	0.008
8	Feature 140	0.382	0.935	Feature 21	0.008
9	Feature 108	0.424	0.884	Feature 6	0.008
10	Feature 10	0.467	0.881	Feature 133	0.008
11	Feature 28	0.502	0.850	Feature 19	0.008
12	Feature 41	0.542	0.848	Feature 26	0.008
13	Feature 18	0.581	0.823	Feature 35	0.008
14	Feature 6	0.615	0.811	Feature 37	0.008
15	Feature 91	0.643	0.745	Feature 38	0.008
16	Feature 89	0.672	0.711	Feature 10	0.008
17	Feature 134	0.685	0.718	Feature 1	0.008
18	Feature 11	0.704	0.612	Feature 34	0.007
19	Feature 32	0.718	0.596	Feature 130	0.007
20	Feature 12	0.734	0.566	Feature 33	0.007

Table 4.20 Top 20 features of the slope model using STEPDISC and BPNN for cereals and buckwheat, oilseeds and pulses

Rank	Selected features for hold-out method	Average squared canonical correlation	Partial r^2	Selected features for BPNN	Weight
1	Feature 294	0.050	0.050	Feature 288	0.010
2	Feature 188	0.099	0.099	Feature 164	0.009
3	Feature 168	0.147	0.147	Feature 165	0.009
4	Feature 206	0.195	0.195	Feature 185	0.009
5	Feature 242	0.243	0.243	Feature 163	0.008
6	Feature 196	0.290	0.290	Feature 197	0.008
7	Feature 238	0.336	0.336	Feature 254	0.008
8	Feature 193	0.382	0.382	Feature 177	0.008
9	Feature 268	0.424	0.424	Feature 162	0.008
10	Feature 296	0.467	0.467	Feature 289	0.008
11	Feature 264	0.502	0.502	Feature 175	0.008
12	Feature 166	0.542	0.542	Feature 182	0.008
13	Feature 184	0.581	0.581	Feature 191	0.008
14	Feature 197	0.615	0.615	Feature 193	0.008
15	Feature 174	0.643	0.643	Feature 194	0.008
16	Feature 162	0.672	0.672	Feature 166	0.008
17	Feature 247	0.685	0.685	Feature 157	0.008
18	Feature 245	0.704	0.704	Feature 190	0.007
19	Feature 290	0.718	0.718	Feature 286	0.007
20	Feature 167	0.734	0.734	Feature 189	0.007

Table 4.21 Top 20 features of the slope-ratio model using STEPDISC and BPNN for cereals and buckwheat, oilseeds and pulses

Rank	Selected features for hold-out method	Average squared canonical correlation	Partial r^2	Selected features for BPNN	Weight
1	Feature 436	0.006	1.000	Feature 375	0.013
2	Feature 412	0.013	0.999	Feature 401	0.012
3	Feature 400	0.019	0.998	Feature 349	0.012
4	Feature 425	0.025	0.997	Feature 330	0.011
5	Feature 413	0.027	0.989	Feature 331	0.011
6	Feature 448	0.033	0.961	Feature 420	0.011
7	Feature 454	0.039	0.952	Feature 423	0.011
8	Feature 342	0.045	0.952	Feature 332	0.011
9	Feature 424	0.051	0.941	Feature 396	0.011
10	Feature 437	0.057	0.939	Feature 351	0.010
11	Feature 373	0.063	0.934	Feature 397	0.010
12	Feature 458	0.068	0.922	Feature 350	0.010
13	Feature 452	0.074	0.903	Feature 347	0.010
14	Feature 379	0.080	0.872	Feature 348	0.010
15	Feature 345	0.085	0.858	Feature 398	0.010
16	Feature 394	0.090	0.797	Feature 402	0.010
17	Feature 395	0.095	0.795	Feature 417	0.010
18	Feature 393	0.099	0.807	Feature 414	0.010
19	Feature 450	0.103	0.794	Feature 344	0.010
20	Feature 315	0.107	0.826	Feature 422	0.009

Table 4.22 Classification accuracies of the combined grains using the ratio model in non- parametric estimation and BPNN

Cereals and buckwheat, oilseeds and pulses	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
2 row barley				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
6 row barley				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Buck wheat				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
CWAD				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
CWRS				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Oats				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Rye				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
SWSW				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Canola				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Flax				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

Table 4.22 continued. Classification accuracies of the combined grains using the ratio model in non- parametric estimation and BPNN

Cereals and buckwheat, oilseeds and pulses	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
Sunflower				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Black beans				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Croma yellow peas				
Set 1	6		6	
Set 2 **	5	94.4$\pm$ 9.6	6	100
Set 3	6		6	
Dark speckled green lentils				
Set 1	6		6	
Set 2 ***	6	94.4$\pm$ 9.6	6	100
Set 3	5		6	
Espace green peas				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Eston lentils				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Laird Lentils				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Light red kidney beans				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Navy pea beans				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.
** Croma yellow peas was misclassified as navy pea beans.
*** Dark speckled green lentils was misclassified as Eston lentils.

Table 4.22 continued. Classification accuracies of the combined grains using the ratio model in non- parametric estimation and BPNN

Cereals and buckwheat, oilseeds and pulses	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
Pinto beans				
Set 1	6		6	
Set 2 ****	5	94.4$\pm$ 9.6	6	100
Set 3	6		6	
Small reds				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.
****Pinto beans was misclassified as navy pea beans.

Table 4.23. Classification accuracies of the combined grains using the slope model in non- parametric estimation and BPNN

Cereals and buckwheat, oilseeds and pulses	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
2 row barley				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
6 row barley				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Buck wheat				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
CWAD				
Set 1	6		6	
Set 2	6	100	6	94.4 ± 9.6
Set 3	6		5	
CWRS				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Oats				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Rye				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
SWSW				
Set 1	4		6	
Set 2	6	83.3 ± 8.5	6	100
Set 3	5		6	
Canola				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Flax				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

Table 4.23 continued. Classification accuracies of the combined grains using the slope model in non- parametric estimation and BPNN

Cereals and buckwheat, oilseeds and pulses	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
Sunflower				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Black beans				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Croma yellow peas				
Set 1	6		5	
Set 2	6	100	5	88.8±9.6
Set 3	6		6	
Dark speckled green lentils				
Set 1	6		5	
Set 2	5	88.8±9.6	5	83.3±8.5
Set 3	5		5	
Espace green peas				
Set 1	6		6	
Set 2	6	100	5	94.4± 9.6
Set 3	6		6	
Eston lentils				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Laird Lentils				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Light red kidney beans				
Set 1	6		6	
Set 2	5	94.4± 9.6	6	100
Set 3	6		6	
Navy pea beans				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

Table 4.23 continued. Classification accuracies of the combined grains using the slope model in non- parametric estimation and BPNN

Cereals and buckwheat, oilseeds and pulses	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
Pinto beans				
Set 1	6		5	
Set 2	5	88.8±9.6	5	88.8±9.6
Set 3	5		6	
Small reds				
Set 1	6		6	
Set 2	6	100	5	94.4± 9.6
Set 3	6		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

Table 4.24. Classification accuracies of the combined grains using the slope-ratio model in non- parametric estimation and BPNN

Cereals and buckwheat, oilseeds and pulses	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
2 row barley				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
6 row barley				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Buck wheat				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
CWAD				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
CWRS				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Oats				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Rye				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
SWSW				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Canola				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Flax				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

Table 4.24 continued. Classification accuracies of the combined grains using the slope-ratio model in non- parametric estimation and BPNN

Cereals and buckwheat, oilseeds and pulses	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
Sunflower				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Black beans				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Croma yellow peas				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Dark speckled green lentils				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Espace green peas				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Eston lentils				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Laird Lentils				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Light red kidney beans				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Navy pea beans				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

Table 4.24 continued. Classification accuracies of the combined grains using the slope-ratio model in non- parametric estimation and BPNN

Cereals and buckwheat, oilseeds and pulses	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
Pinto beans				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Small reds				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

4.5 Specialty Seeds

From the reflectance data (Fig. A4a, A4b, A4c, A4d and A4e; Appendix A) ratio, slope and slope-ratio features were extracted. The contributions of the individual features to the classifiers were determined and the features were ranked in descending order (Tables 4.25, 4.26, and 4.27), respectively, for the three models. Using the top 20 features for each model classifications were done for the specialty seeds and results are given in Tables 4.28, 4.29, and 4.30, respectively, for the ratio, the slope and, the slope ratio models.

A classification accuracy of 100 % was obtained for the twenty-five specialty seeds using the ratio and the slope-ratio models for both the statistical and the neural network classifier (Tables 4.28 and 4.30). The slope model gave a misclassification for both the BPNN and the non-parametric estimation misclassifying tall fescue as orchard grass when using the non-parametric classifier and tall fescue as slender wheat grass and annual rye grass when the BPNN classifier was used (Table 4.29).

The reflectance characteristics could classify the specialty seeds when used with either the ratio model or the slope ratio model but the slope-ratio model proved to be again a more successful classifier based on the partial r^2 values (Tables 4.25, 4.26, and 4.27). The chosen features were more concentrated in the UV region and the visible region in the ratio and the slope models. The top 20 features in the slope-ratio model were more concentrated in the visible and the near-infrared region. The overlapping of the reflectance characteristics (Fig. A4a, A4b, A4c, A4d and A4e; Appendix A) reflect the negative classification of the slope model.

Here again the possibility of having filters and implementing the process in an industrial setting would be questionable as the top 20 features are spread across the whole range (320 – 1880 nm) and have no similarities with the features in the other sections.

Though the features are different in both the classifiers in all three models as the classifiers function differently to classify, it is interesting to note that the features used for the classification are concentrated in the same region of the electromagnetic spectrum (Tables 4.25, 4.26, and 4.27). This means a classifier could be developed for classification of specialty seeds using the visible spectrum for which design of a spectrophotometer is less complex and thus less expensive.

Table 4.25 Top 20 features of the ratio model using STEPDISC and BPNN for specialty seeds

Rank	Selected features for hold-out method	Average squared canonical correlation	Partial r^2	Selected features for BPNN	Weight
1	Feature 91	0.050	0.991	Feature 7	0.009
2	Feature 86	0.099	0.981	Feature 8	0.009
3	Feature 13	0.147	0.973	Feature 11	0.009
4	Feature 35	0.195	0.968	Feature 12	0.009
5	Feature 40	0.243	0.967	Feature 6	0.009
6	Feature 2	0.291	0.952	Feature 4	0.009
7	Feature 83	0.337	0.942	Feature 72	0.009
8	Feature 24	0.383	0.925	Feature 19	0.008
9	Feature 111	0.427	0.905	Feature 66	0.008
10	Feature 29	0.467	0.888	Feature 18	0.008
11	Feature 89	0.511	0.893	Feature 9	0.008
12	Feature 44	0.547	0.860	Feature 10	0.008
13	Feature 107	0.587	0.855	Feature 55	0.008
14	Feature 102	0.614	0.782	Feature 17	0.008
15	Feature 11	0.645	0.781	Feature 44	0.008
16	Feature 6	0.681	0.761	Feature 33	0.008
17	Feature 69	0.709	0.723	Feature 46	0.008
18	Feature 19	0.735	0.653	Feature 16	0.009
19	Feature 65	0.754	0.575	Feature 3	0.009
20	Feature 37	0.763	0.555	Feature 67	0.009

Table 4.26 Top 20 features of the slope model using STEPDISC and BPNN specialty seeds

Rank	Selected features for hold-out method	Average squared canonical correlation	Partial r^2	Selected features for BPNN	Weight
1	Feature 247	0.050	0.991	Feature 163	0.010
2	Feature 242	0.099	0.981	Feature 164	0.010
3	Feature 169	0.147	0.973	Feature 166	0.009
4	Feature 191	0.195	0.968	Feature 167	0.009
5	Feature 196	0.243	0.967	Feature 162	0.008
6	Feature 158	0.291	0.952	Feature 174	0.008
7	Feature 239	0.337	0.942	Feature 191	0.008
8	Feature 180	0.383	0.925	Feature 160	0.008
9	Feature 267	0.427	0.905	Feature 204	0.008
10	Feature 185	0.467	0.888	Feature 159	0.008
11	Feature 245	0.511	0.893	Feature 165	0.008
12	Feature 200	0.547	0.860	Feature 228	0.008
13	Feature 263	0.587	0.855	Feature 192	0.008
14	Feature 258	0.614	0.782	Feature 173	0.008
15	Feature 167	0.645	0.781	Feature 168	0.008
16	Feature 162	0.681	0.761	Feature 202	0.008
17	Feature 225	0.709	0.723	Feature 232	0.008
18	Feature 175	0.735	0.653	Feature 175	0.008
19	Feature 221	0.754	0.575	Feature 206	0.007
20	Feature 193	0.763	0.555	Feature 208	0.007

Table 4.27 Top 20 features of the slope-ratio model using STEPDISC and BPNN specialty seeds

Rank	Selected features for hold-out method	Average squared canonical correlation	Partial r^2	Selected features for BPNN	Weight
1	Feature 403	0.050	0.991	Feature 323	0.013
2	Feature 398	0.099	0.981	Feature 330	0.013
3	Feature 325	0.147	0.973	Feature 337	0.013
4	Feature 347	0.195	0.968	Feature 348	0.013
5	Feature 352	0.243	0.967	Feature 341	0.013
6	Feature 314	0.291	0.952	Feature 400	0.013
7	Feature 395	0.337	0.942	Feature 339	0.013
8	Feature 336	0.383	0.925	Feature 342	0.012
9	Feature 423	0.427	0.905	Feature 347	0.012
10	Feature 341	0.467	0.888	Feature 414	0.012
11	Feature 401	0.511	0.893	Feature 401	0.012
12	Feature 356	0.547	0.860	Feature 350	0.012
13	Feature 419	0.587	0.855	Feature 336	0.011
14	Feature 414	0.614	0.782	Feature 349	0.011
15	Feature 323	0.645	0.781	Feature 420	0.011
16	Feature 318	0.681	0.761	Feature 421	0.011
17	Feature 381	0.709	0.723	Feature 449	0.011
18	Feature 331	0.735	0.653	Feature 322	0.010
19	Feature 377	0.754	0.575	Feature 418	0.010
20	Feature 349	0.763	0.555	Feature 351	0.010

Table 4.28. Classification accuracies of the specialty seeds using the ratio model in non- parametric estimation and BPNN

Specialty Seeds	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
Alfalfa				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Alsike Clover				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Annual rye grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Birds foot trefoil				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Brown mustard				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Creeping bent grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Creeping red fescue				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Crested wheat grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Crown millet				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Intermediate wheat grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

Table 4.28 continued. Classification accuracies of the specialty seeds using the ratio model in non- parametric estimation and BPNN

Specialty Seeds	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
Kentucky blue grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Meadow brome grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Meadow fescue				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Orchard grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Oriental mustard				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Perennial rye grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Red clover				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Reed canary grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Siberian millet				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

Table 4.28 continued. Classification accuracies of the specialty seeds using the ratio model in non- parametric estimation and BPNN

Specialty Seeds	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
Slender wheat grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Sorghum Sudan grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Sweet clover				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Tall fescue				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Timothy				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Yellow mustard				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

Table 4.29. Classification accuracies of the specialty seeds using the slope model in non- parametric estimation and BPNN

Specialty seeds	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
Alfalfa				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Alsike Clover				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Annual rye grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Birds foot trefoil				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Brown mustard				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Creeping bent grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Creeping red fescue				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Crested wheat grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Crown millet				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Intermediate wheat grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

Table 4.29 continued. Classification accuracies of the specialty seeds using the slope model in non- parametric estimation and BPNN

Specialty Seeds	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
Kentucky blue grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Meadow brome grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Meadow fescue				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Orchard grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Oriental mustard				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Perennial rye grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Red clover				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Reed canary grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Siberian millet				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

Table 4.29 continued. Classification accuracies the specialty seeds using the slope model in non- parametric estimation and BPNN

Specialty Seeds	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
Slender wheat grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Sorghum Sudan grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Sweet clover				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Tall fescue				
Set 1	5		5	
Set 2**	6	94.4±9.6	6	88.8±9.6
Set 3	6		5	
Timothy				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Yellow mustard				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

** Tall fescue was misclassified as Orchard grass when non-parametric estimation was used, and was misclassified once as Annual rye grass and once as slender wheat grass when BPNN classifier was used.

Table 4.30. Classification accuracies of the specialty seeds using the slope-ratio model in non- parametric estimation and BPNN

Specialty seeds	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
Alfalfa				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Alsike Clover				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Annual rye grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Birds foot trefoil				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Brown mustard				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Creeping bent grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Creeping red fescue				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Crested wheat grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Crown millet				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Intermediate wheat grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

Table 4.30 continued. Classification accuracies of the specialty seeds using the slope-ratio in non- parametric estimation and BPNN

Specialty Seeds	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
Kentucky blue grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Meadow brome grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Meadow fescue				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Orchard grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Oriental mustard				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Perennial rye grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Red clover				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Reed canary grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Siberian millet				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

Table 4.30 continued. Classification accuracies of the specialty seeds using the slope-ratio model in non- parametric estimation and BPNN

Specialty Seeds	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
Slender wheat grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Sorghum Sudan grass				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Sweet clover				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Tall fescue				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Timothy				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
Yellow mustard				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

4.6 Canada Western Red Spring (CWRS) Wheat Grading Characteristics

Reflectance characteristics were not capable of classifying differences among the growing regions and crop-year, moisture content, or foreign material content with 100 % accuracy. The results imply that the variations in the CWRS wheat samples from across the prairies are similar to variations from a single location. Different moisture contents and presence of foreign material has more effect on the reflectance characteristics than the growing regions.

4.6.1 Growing Region and Crop-Year

From the reflectance data (Fig. A5a, A5b, and A5c; Appendix A) ratio, slope and slope-ratio features were extracted. The contributions of the individual features to the classifiers were determined and the features were ranked in descending order (Tables 4.31, 4.32, and 4.33). Using the top 20 features for each model classifications were done for CWRS wheat based on the growing regions and results are given in Tables 4.34, 4.35, and 4.36, respectively, for the ratio, the slope and, the slope ratio models. The confusion matrices for the classification accuracies are presented in Tables B2a, B2b, B2c, B2d, B2e, and B2f, for the three models and two classifiers (statistical and BPNN), respectively.

The reflectance curves, given in figures A5a, A5b, and A5c; Appendix A show that there was hardly any difference among the reflectance data collected from the same sample 30 times (Fig. 2.3.). Another observation that can be seen is that the features (Tables 4.31, 4.32, and 4.33) chosen by the classifiers for their classification processes lie more in the visible and the near-infrared region which have the reflectance data more spread out for each growing region than the UV region where the reflectance data collected for the different growing regions and crop-years are more clustered and have very negligible difference.

The confusion matrices (Tables B2a, B2b, B2c, B2d, B2e, and B2f. Appendix B) show the inaccurate classification of the indicating that the growing regions have no effect on the reflectance characteristics. The classification accuracies ranged from a low 0 % to a high of 77.7 % showing that reflectance characteristics could be used in the automation process without bias by the growing region.

Table 4.31 Top 20 features of the ratio model using STEPDISC and BPNN for different growing regions and crop-years of CWRS wheat

Rank	Selected features for hold-out method	Average squared canonical correlation	Partial r^2	Selected features for BPNN	Weight
1	Feature 116	0.009	0.743	Feature 47	0.009
2	Feature 143	0.018	0.715	Feature 121	0.009
3	Feature 120	0.023	0.578	Feature 99	0.009
4	Feature 112	0.030	0.570	Feature 65	0.009
5	Feature 50	0.037	0.506	Feature 97	0.009
6	Feature 140	0.043	0.506	Feature 62	0.009
7	Feature 34	0.049	0.489	Feature 78	0.009
8	Feature 125	0.054	0.569	Feature 154	0.008
9	Feature 102	0.060	0.502	Feature 137	0.008
10	Feature 27	0.065	0.493	Feature 144	0.008
11	Feature 21	0.070	0.486	Feature 142	0.008
12	Feature 66	0.075	0.470	Feature 56	0.008
13	Feature 62	0.081	0.485	Feature 44	0.008
14	Feature 113	0.086	0.477	Feature 95	0.008
15	Feature 71	0.091	0.455	Feature 98	0.008
16	Feature 40	0.097	0.451	Feature 107	0.008
17	Feature 38	0.102	0.449	Feature 30	0.008
18	Feature 63	0.107	0.448	Feature 55	0.007
19	Feature 84	0.112	0.474	Feature 135	0.007
20	Feature 146	0.117	0.460	Feature 103	0.007

Table 4.32 Top 20 features of the slope model using STEPDISC and BPNN for different growing regions and crop-years of CWRS wheat

Rank	Selected features for hold-out method	Average squared canonical correlation	Partial r^2	Selected features for BPNN	Weight
1	Feature 254	0.056	0.787	Feature 244	0.009
2	Feature 256	0.103	0.701	Feature 310	0.009
3	Feature 244	0.120	0.624	Feature 234	0.008
4	Feature 253	0.147	0.496	Feature 255	0.008
5	Feature 255	0.154	0.343	Feature 254	0.008
6	Feature 271	0.162	0.328	Feature 220	0.008
7	Feature 221	0.174	0.310	Feature 222	0.008
8	Feature 224	0.188	0.296	Feature 223	0.008
9	Feature 200	0.202	0.314	Feature 259	0.008
10	Feature 193	0.214	0.272	Feature 226	0.008
11	Feature 310	0.230	0.261	Feature 245	0.008
12	Feature 212	0.246	0.259	Feature 203	0.008
13	Feature 163	0.262	0.265	Feature 218	0.008
14	Feature 220	0.267	0.214	Feature 300	0.008
15	Feature 168	0.278	0.207	Feature 232	0.008
16	Feature 300	0.287	0.218	Feature 210	0.008
17	Feature 209	0.297	0.214	Feature 221	0.008
18	Feature 238	0.301	0.210	Feature 216	0.008
19	Feature 217	0.311	0.191	Feature 311	0.008
20	Feature 304	0.317	0.172	Feature 309	0.008

Table 4.33 Top 20 features of the slope-ratio model using STEPDISC and BPNN for different growing regions and crop-years of CWRS wheat

Rank	Selected features for hold-out method	Average squared canonical correlation	Partial r^2	Selected features for BPNN	Weight
1	Feature 354	0.009	0.743	Feature 434	0.008
2	Feature 356	0.018	0.715	Feature 410	0.008
3	Feature 344	0.023	0.578	Feature 384	0.008
4	Feature 353	0.030	0.570	Feature 385	0.008
5	Feature 455	0.037	0.506	Feature 394	0.008
6	Feature 371	0.043	0.506	Feature 420	0.008
7	Feature 421	0.049	0.489	Feature 352	0.008
8	Feature 424	0.054	0.569	Feature 333	0.008
9	Feature 400	0.060	0.502	Feature 349	0.007
10	Feature 383	0.065	0.493	Feature 426	0.008
11	Feature 410	0.070	0.486	Feature 445	0.008
12	Feature 313	0.075	0.470	Feature 403	0.008
13	Feature 363	0.081	0.485	Feature 418	0.008
14	Feature 320	0.086	0.477	Feature 400	0.008
15	Feature 322	0.091	0.455	Feature 432	0.008
16	Feature 398	0.097	0.451	Feature 410	0.008
17	Feature 409	0.102	0.449	Feature 321	0.007
18	Feature 438	0.107	0.448	Feature 216	0.007
19	Feature 417	0.112	0.474	Feature 411	0.007
20	Feature 404	0.117	0.460	Feature 412	0.007

Table 4.34. Classification accuracies for the growing regions and crop years of CWRS wheat using the ratio model in non- parametric estimation and BPNN

Growing regions	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
Burdett				
Set 1	1		1	
Set 2	2	16.6±16.6	1	11.1±19.2
Set 3	0		0	
Camrose				
Set 1	0		0	
Set 2	0	0	1	5.5±3.2
Set 3	0		0	
Churchill				
Set 1	2		1	
Set 2	0	11.1±19.2	1	11.1±19.2
Set 3	0		0	
Estevan				
Set 1	0		0	
Set 2	1	16.6±16.6	1	16.6±16.6
Set 3	2		2	
Fair view				
Set 1	2		2	
Set 2	0	16.6±16.6	0	16.6±16.6
Set 3	1		1	
Lloyd Minster				
Set 1	0		1	
Set 2	0	0	1	11.1±19.2
Set 3	0		0	
Medicine hat				
Set 1	1		1	
Set 2	0	5.5±3.2	0	5.5±3.2
Set 3	0		0	
Melita				
Set 1	0		0	
Set 2	0	0	0	0
Set 3	0		0	
North Battleford				
Set 1	2		3	
Set 2	2	38.8±9.6	4	50±25.46
Set 3	3		1	
Prince Albert				
Set 1	0		0	
Set 2	3	16.6±16.6	1	5.5±3.2
Set 3	0		0	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

Table 4.34 continued. Classification accuracies for the growing regions and crop-years of CWRS wheat using the ratio model in non- parametric estimation and BPNN

Growing Regions	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
Saskatoon				
Set 1	1		1	
Set 2	3	33.3±16.7	0	11.1±19.2
Set 3	2		1	
Swift current				
Set 1	1		1	
Set 2	0	5.5±3.2	0	5.5±3.2
Set 3	0		0	
Tisdale				
Set 1	1		0	
Set 2	0	11.1±19.2	0	5.5±3.2
Set 3	1		1	
Vegerville				
Set 1	0		0	
Set 2	2	16.6±16.6	2	16.6±16.6
Set 3	1		1	
Vermilion				
Set 1	0		0	
Set 2	2	27.7±25.45	1	11.1±19.2
Set 3	3		1	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

Table 4.35. Classification accuracies for the growing regions and crop-years of CWRS wheat using the slope model in non- parametric estimation and BPNN

Growing regions	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
Burdett				
Set 1	1		1	
Set 2	1	16.6	0	0.05±9.58
Set 3	1		0	
Camrose				
Set 1	1		2	
Set 2	0	22.2±25.4	2	44.4±19.22
Set 3	3		4	
Churchill				
Set 1	2		4	
Set 2	0	11.1±19.2	5	55.5±25.5
Set 3	0		2	
Estevan				
Set 1	2		3	
Set 2	1	27.7±25.45	1	27.7±19.3
Set 3	2		1	
Fair view				
Set 1	1		2	
Set 2	0	11.1±19.2	0	22.2±19.2
Set 3	1		2	
Lloyd Minster				
Set 1	1		2	
Set 2	1	16.6	0	16.6±16.7
Set 3	1		1	
Medicine hat				
Set 1	0		0	
Set 2	1	5.5±3.2	0	0
Set 3	0		0	
Melita				
Set 1	2		3	
Set 2	0	11.1±19.2	4	55.5±9.6
Set 3	0		3	
North Battleford				
Set 1	3		1	
Set 2	5	61.1±25.4	3	22.2±25.5
Set 3	2		0	
Prince Albert				
Set 1	3		1	
Set 2	0	33.33±28.86	1	22.2±9.6
Set 3	3		2	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

Table 4.35 continued. Classification accuracies for the growing regions and crop-years of CWRS wheat using the slope model in non- parametric estimation and BPNN

Growing Regions	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
Saskatoon				
Set 1	3		1	
Set 2	1	33.3±16.7	4	27.7±34.7
Set 3	2		0	
Swift current				
Set 1	1		0	
Set 2	1	16.6	0	0
Set 3	1		0	
Tisdale				
Set 1	1		0	
Set 2	1	16.6	0	0
Set 3	1		0	
Vegerville				
Set 1	2		0	
Set 2	0	22.2±19.22	0	0
Set 3	2		0	
Vermilion				
Set 1	0		0	
Set 2	2	33.3±34.66	0	0
Set 3	4		0	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

Table 4.36. Classification accuracies for the growing regions and crop-years of CWRS wheat using the slope-ratio model in non- parametric estimation and BPNN

Growing regions	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
Burdett				
Set 1	2		2	
Set 2	2	22.2±19.2	4	33.3±33.3
Set 3	0		0	
Camrose				
Set 1	1		4	
Set 2	2	22.2±9.7	5	55.5±25.5
Set 3	1		1	
Churchill				
Set 1	4		3	
Set 2	3	55.5±9.6	5	55.5±25.5
Set 3	3		2	
Estevan				
Set 1	0		1	
Set 2	0	0	3	27.7±19.28
Set 3	0		1	
Fair view				
Set 1	3		2	
Set 2	0	33.2±25.5	2	33.3
Set 3	1		2	
Lloyd Minster				
Set 1	3		2	
Set 2	3	50	3	38.8±9.6
Set 3	3		2	
Medicine hat				
Set 1	3		2	
Set 2	1	33.3±8.4	1	22.2±9.7
Set 3	2		1	
Melita				
Set 1	1		2	
Set 2	1	16.6	1	22.2±9.7
Set 3	1		1	
North Battleford				
Set 1	4		6	
Set 2	6	77.7±19.3	3	72.2±25.5
Set 3	4		4	
Prince Albert				
Set 1	1		1	
Set 2	3	27.7±19.28	1	16.6
Set 3	1		1	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

Table 4.36 continued. Classification accuracies for the growing regions and crop-years of CWRS wheat using the slope-ratio model in non- parametric estimation and BPNN

Growing regions	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
Saskatoon				
Set 1	1		2	
Set 2	1	22.2±9.7	3	38.8±9.6
Set 3	2		2	
Swift current				
Set 1	1		4	
Set 2	4	27.7±34.7	1	27.7±34.7
Set 3	0		0	
Tisdale				
Set 1	2		0	
Set 2	0	11.1±25.1	0	22.2±38.5
Set 3	3		4	
Vegerville				
Set 1	2		0	
Set 2	5	44.4±36.9	3	33.3±28.8
Set 3	1		3	
Vermilion				
Set 1	4		1	
Set 2	2	50±16.6	5	50±33.3
Set 3	3		3	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

4.6.2 Moisture Content

From the reflectance data (Fig. A7. Appendix A) ratio, slope and slope-ratio features were extracted. The contributions of the individual features to the classifiers were determined and the features were ranked in descending order (Tables 4.37, 4.38, and 4.39). Using the top 20 features for each model classifications were done for CWRS wheat based on the five different moisture contents, namely 10, 12, 14, 16, and 18 % and results are given in Tables 4.40, 4.41, and 4.42, respectively, for the ratio, slope and, slope ratio models. The confusion matrices for the classification accuracies are in Tables B3a, B3b, B3c, B3d, B3e, and B3f, Appendix B, for the three models, respectively.

The accuracies ranged from a high of 100 % accuracy for 10 % mc in the ratio model for both the BPNN and non-parametric estimation to a low of 44.4 % accuracy for BPNN in the slope model. The ratio model gave better classification in both the BPNN and the non-parametric estimation than the slope model or the slope-ratio model.

The classification accuracies show that the reflectance characteristics could not correctly classify the grain at a 100 % accuracy, based on the moisture contents. However it can be seen that the misclassification was between adjacent moisture content samples (Tables B3a, B3b, B3c, B3d, B3e, and B3f. Appendix B), i.e., samples differing by 2 % moisture content could be differentiated.

Table 4.37 Top 20 features of the ratio model using STEPDISC and BPNN for different moisture contents of CWRS wheat

Rank	Selected features for hold-out method	Average squared canonical correlation	Partial r^2	Selected features for BPNN	Weight
1	Feature 66	0.138	0.968	Feature 104	0.007
2	Feature 13	0.274	0.957	Feature 143	0.007
3	Feature 111	0.406	0.921	Feature 74	0.007
4	Feature 29	0.525	0.879	Feature 10	0.007
5	Feature 19	0.626	0.779	Feature 71	0.007
6	Feature 17	0.726	0.792	Feature 105	0.007
7	Feature 6	0.818	0.722	Feature 139	0.007
8	Feature 137	0.838	0.676	Feature 117	0.007
9	Feature 21	0.855	0.427	Feature 63	0.007
10	Feature 89	0.876	0.412	Feature 21	0.007
11	Feature 156	0.882	0.375	Feature 22	0.007
12	Feature 45	0.892	0.373	Feature 85	0.007
13	Feature 10	0.901	0.361	Feature 118	0.007
14	Feature 57	0.904	0.344	Feature 29	0.007
15	Feature 36	0.909	0.328	Feature 54	0.007
16	Feature 97	0.912	0.283	Feature 136	0.007
17	Feature 56	0.919	0.284	Feature 127	0.007
18	Feature 133	0.921	0.260	Feature 134	0.007
19	Feature 99	0.923	0.247	Feature 120	0.007
20	Feature 9	0.926	0.200	Feature 43	0.007

Table 4.38 Top 20 features of the slope model using STEPDISC and BPNN for different moisture contents of CWRS wheat

Rank	Selected features for hold-out method	Average squared canonical correlation	Partial r^2	Selected features for BPNN	Weight
1	Feature 196	0.138	0.968	Feature 159	0.009
2	Feature 169	0.274	0.957	Feature 158	0.008
3	Feature 267	0.406	0.921	Feature 227	0.008
4	Feature 185	0.525	0.879	Feature 219	0.008
5	Feature 175	0.626	0.779	Feature 273	0.008
6	Feature 173	0.726	0.792	Feature 230	0.008
7	Feature 162	0.818	0.722	Feature 260	0.008
8	Feature 293	0.838	0.676	Feature 299	0.007
9	Feature 177	0.855	0.427	Feature 166	0.007
10	Feature 245	0.876	0.412	Feature 295	0.007
11	Feature 311	0.882	0.375	Feature 261	0.007
12	Feature 201	0.892	0.373	Feature 160	0.007
13	Feature 166	0.901	0.361	Feature 269	0.007
14	Feature 213	0.904	0.344	Feature 292	0.007
15	Feature 192	0.909	0.328	Feature 187	0.007
16	Feature 253	0.912	0.283	Feature 274	0.007
17	Feature 212	0.919	0.284	Feature 177	0.007
18	Feature 289	0.921	0.260	Feature 290	0.007
19	Feature 255	0.923	0.247	Feature 205	0.007
20	Feature 165	0.926	0.200	Feature 192	0.007

Table 4.39 Top 20 features of the slope-ratio model using STEPDISC and BPNN for different moisture contents of CWRS wheat

Rank	Selected features for hold-out method	Average squared canonical correlation	Partial r^2	Selected features for BPNN	Weight
1	Feature 465	0.022	0.999	Feature 418	0.007
2	Feature 391	0.043	0.993	Feature 399	0.007
3	Feature 359	0.064	0.970	Feature 339	0.007
4	Feature 364	0.085	0.972	Feature 393	0.007
5	Feature 372	0.106	0.972	Feature 329	0.007
6	Feature 365	0.126	0.935	Feature 381	0.007
7	Feature 452	0.140	0.933	Feature 461	0.007
8	Feature 374	0.158	0.938	Feature 322	0.007
9	Feature 446	0.175	0.934	Feature 432	0.007
10	Feature 456	0.190	0.920	Feature 405	0.007
11	Feature 358	0.205	0.895	Feature 388	0.007
12	Feature 450	0.221	0.914	Feature 321	0.007
13	Feature 315	0.234	0.902	Feature 334	0.007
14	Feature 338	0.247	0.923	Feature 451	0.007
15	Feature 332	0.261	0.918	Feature 397	0.007
16	Feature 343	0.277	0.922	Feature 419	0.007
17	Feature 346	0.288	0.943	Feature 427	0.007
18	Feature 436	0.306	0.919	Feature 366	0.007
19	Feature 434	0.321	0.929	Feature 454	0.007
20	Feature 352	0.335	0.942	Feature 402	0.007

Table 4.40. Classification accuracies for different moisture contents of CWRS wheat using the ratio model in non- parametric estimation and BPNN

Moisture content	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
MC 10				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
MC12				
Set 1	4		4	
Set 2	6	83.3±16.7	5	83.3±16.7
Set 3	5		6	
MC14				
Set 1	5		5	
Set 2	5	88.8±9.6	5	88.8±9.6
Set 3	6		6	
MC 16				
Set 1	5		5	
Set 2	6	88.8±9.6	6	94.4±9.6
Set 3	5		6	
MC 18				
Set 1	6		5	
Set 2	6	100	6	94.4±9.6
Set 3	6		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

Table 4.41. Classification accuracies for different moisture contents of CWRS wheat using the slope model in non- parametric estimation and BPNN

Moisture content	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
MC 10				
Set 1	4		3	
Set 2	4	77.7±19.3	6	77.7±25.5
Set 3	6		5	
MC12				
Set 1	2		4	
Set 2	4	83.3±16.7	2	83.3±16.7
Set 3	3		3	
MC14				
Set 1	5		2	
Set 2	2	61.1±25.5	3	44.4±19.3
Set 3	4		3	
MC 16				
Set 1	5		4	
Set 2	6	94.4±9.6	3	83.3±16.7
Set 3	6		2	
MC 18				
Set 1	5		3	
Set 2	6	94.4±9.6	5	77.7±25.5
Set 3	6		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

Table 4.42. Classification accuracies for different moisture contents of CWRS wheat using the slope-ratio model in non- parametric estimation and BPNN

Moisture content	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
MC 10				
Set 1	5		5	
Set 2	6	94.4±9.6	6	94.4±9.6
Set 3	6		6	
MC12				
Set 1	4		4	
Set 2	4	77.7±19.3	6	83.3±16.7
Set 3	6		5	
MC14				
Set 1	4		6	
Set 2	5	83.3±16.7	3	83.3±28.9
Set 3	6		6	
MC 16				
Set 1	5		4	
Set 2	3	83.3±25.5	6	83.3±16.7
Set 3	6		5	
MC 18				
Set 1	5		6	
Set 2	6	88.8±9.6	6	100
Set 3	5		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

4.6.3 Foreign Material Content

From the reflectance data of CWRS wheat with different foreign material contents (Fig. A8. Appendix A) ratio, slope, and slope-ratio features were extracted. The contributions of the individual features to the classifiers were determined and the features were ranked in descending order (Tables 4.43, 4.44, and 4.45). Using the top 20 features for each model, classifications were done for CWRS wheat based on the different compositions of foreign material namely, 3, 6, 9, 12, 15 % foreign material content and results are given in Tables 4.34, 4.35 and, 4.36, respectively for the ratio, the slope and, the slope ratio models. The confusion matrices for classification accuracies are in Tables B4a, B4b, B4c, B4d, B4e, and B4f, Appendix B, for the three models respectively.

The reflectance data in Fig. A8 show visible difference from the thirty replicates of the same CWRS wheat sample shown in Fig 2.3. The features selected by the classifiers covered the whole scanned spectrum.

The slope-ratio model gave the highest classification accuracy followed by the ratio model and the slope model. The BPNN classifier gave a higher accuracy than the non-parametric classifier by only misclassifying 12 % foreign material wheat. The use of all the top 20 features to get this accuracy has eliminated the possibility of introducing filters and performing the classification in an automated simplified process.

The errors in classification could have been due to the sample preparation where the concentration of the foreign material could have been in any one part of the sample. The other possibility would be that the foreign material was not exposed to the incident light, as there would be only one surface of the sample holder facing the beam. Because of these possible sources of errors, it would be a difficult task to use reflectance characteristics as a positive identifier for the amount of foreign material in a given sample of wheat.

Table 4.43 Top 20 features of the ratio model using STEPDISC and BPNN for different foreign materials in CWRS wheat samples

Rank	Selected features for hold-out method	Average squared canonical correlation	Partial r^2	Selected features for BPNN	Weight
1	Feature 66	0.138	0.968	Feature 102	0.008
2	Feature 13	0.274	0.957	Feature 152	0.007
3	Feature 111	0.406	0.921	Feature 140	0.007
4	Feature 29	0.525	0.879	Feature 34	0.007
5	Feature 19	0.626	0.779	Feature 10	0.007
6	Feature 17	0.726	0.792	Feature 26	0.007
7	Feature 6	0.818	0.722	Feature 136	0.007
8	Feature 137	0.838	0.676	Feature 21	0.007
9	Feature 21	0.855	0.427	Feature 16	0.007
10	Feature 89	0.876	0.412	Feature 131	0.007
11	Feature 156	0.882	0.375	Feature 134	0.007
12	Feature 45	0.892	0.373	Feature 103	0.007
13	Feature 10	0.901	0.361	Feature 8	0.007
14	Feature 57	0.904	0.344	Feature 62	0.007
15	Feature 36	0.909	0.328	Feature 86	0.007
16	Feature 97	0.912	0.283	Feature 110	0.007
17	Feature 56	0.919	0.284	Feature 117	0.007
18	Feature 133	0.921	0.260	Feature 19	0.007
19	Feature 99	0.923	0.247	Feature 115	0.007
20	Feature 9	0.926	0.200	Feature 127	0.007

Table 4.44 Top 20 features of the slope model using STEPDISC and BPNN for different foreign materials in CWRS wheat samples

Rank	Selected features for hold-out method	Average squared canonical correlatio	Partial r^2	Selected features for BPNN	Weight
1	Feature 222	0.138	0.968	Feature 159	0.008
2	Feature 169	0.274	0.957	Feature 158	0.007
3	Feature 267	0.406	0.921	Feature 227	0.007
4	Feature 185	0.525	0.879	Feature 219	0.007
5	Feature 175	0.626	0.779	Feature 273	0.007
6	Feature 173	0.726	0.792	Feature 230	0.007
7	Feature 162	0.818	0.722	Feature 260	0.007
8	Feature 293	0.838	0.676	Feature 299	0.007
9	Feature 177	0.855	0.427	Feature 166	0.007
10	Feature 245	0.876	0.412	Feature 295	0.007
11	Feature 310	0.882	0.375	Feature 261	0.007
12	Feature 201	0.892	0.373	Feature 160	0.007
13	Feature 166	0.901	0.361	Feature 269	0.007
14	Feature 213	0.904	0.344	Feature 292	0.007
15	Feature 192	0.909	0.328	Feature 187	0.007
16	Feature 253	0.912	0.283	Feature 274	0.007
17	Feature 212	0.919	0.284	Feature 177	0.007
18	Feature 289	0.921	0.260	Feature 290	0.007
19	Feature 254	0.923	0.247	Feature 205	0.007
20	Feature 165	0.926	0.200	Feature 192	0.007

Table 4.45 Top 20 features of the slope-ratio model using STEPDISC and BPNN for different foreign materials in CWRS wheat samples

Rank	Selected features for hold-out method	Average squared canonical correlation	Partial r^2	Selected features for BPNN	Weight
1	Feature 315	0.029	0.999	Feature 414	0.008
2	Feature 455	0.057	0.989	Feature 464	0.007
3	Feature 456	0.086	0.980	Feature 452	0.007
4	Feature 374	0.114	0.984	Feature 346	0.007
5	Feature 366	0.136	0.986	Feature 322	0.007
6	Feature 453	0.164	0.990	Feature 338	0.007
7	Feature 380	0.186	0.991	Feature 448	0.007
8	Feature 350	0.213	0.995	Feature 333	0.007
9	Feature 358	0.236	1.000	Feature 328	0.007
10	Feature 451	0.253	1.000	Feature 443	0.007
11	Feature 343	0.275	1.000	Feature 446	0.007
12	Feature 408	0.295	1.000	Feature 415	0.007
13	Feature 411	0.310	1.000	Feature 320	0.007
14	Feature 391	0.320	1.000	Feature 374	0.007
15	Feature 433	0.343	1.000	Feature 398	0.007
16	Feature 439	0.360	1.000	Feature 422	0.007
17	Feature 444	0.377	1.000	Feature 429	0.007
18	Feature 465	0.394	1.000	Feature 331	0.007
19	Feature 452	0.410	1.000	Feature 427	0.007
20	Feature 361	0.431	1.000	Feature 439	0.007

Table 4.46. Classification accuracies for different foreign material content CWRS wheat samples using the ratio model in non- parametric estimation and BPNN

Foreign Material	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
FM 03				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
FM06				
Set 1	6		5	
Set 2	5	88.8±9.6	6	83.3±9.6
Set 3	5		5	
FM09				
Set 1	5		5	
Set 2	5	77.7±9.6	5	77.7±9.6
Set 3	4		4	
FM 12				
Set 1	3		4	
Set 2	5	72.2±19.3	5	77.7±9.6
Set 3	5		5	
FM 15				
Set 1	5		6	
Set 2	6	94.4±9.6	6	100
Set 3	6		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

Table 4.47. Classification accuracies for different foreign material content CWRS wheat samples using the slope model in non- parametric estimation and BPNN

Foreign Material	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
FM 03				
Set 1	5		6	
Set 2	5	88.8±9.6	6	100
Set 3	6		6	
FM06				
Set 1	6		5	
Set 2	6	94.4±9.6	5	88.8±9.6
Set 3	5		6	
FM09				
Set 1	6		5	
Set 2	6	88.8±19.3	4	77.7±9.6
Set 3	4		5	
FM 12				
Set 1	6		4	
Set 2	6	100	5	83.3±16.7
Set 3	6		6	
FM 15				
Set 1	5		5	
Set 2	6	94.4±9.6	5	88.8±9.6
Set 3	6		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

Table 4.48. Classification accuracies for different foreign material content CWRS wheat samples using the slope-ratio model in non- parametric estimation and BPNN

Foreign Material	Non-parametric estimation		BPNN	
	No.	Mean (%)*	No.	Mean (%)*
FM 03				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	
FM06				
Set 1	6		6	
Set 2	5	88.8±9.6	6	100
Set 3	5		6	
FM09				
Set 1	6		6	
Set 2	6	94.4±9.6	6	100
Set 3	5		6	
FM 12				
Set 1	5		5	
Set 2	6	94.4±9.6	6	94.4±9.6
Set 3	6		6	
FM 15				
Set 1	6		6	
Set 2	6	100	6	100
Set 3	6		6	

* Mean followed by standard deviation. If standard deviation is not given then it is 0.

5. CONCLUSIONS

The research in this thesis was a part of a major project on the development of a machine vision system for automated classification, sorting and grading of bulk grains at terminal elevators. Reflectance characteristics were measured for seven cereals and buckwheat, three oilseeds, 10 pulses, 25 specialty seeds and for CRWS from different growing regions and crop-years, different moisture contents, and several foreign material contents. A total of 465 features consisting of 156 for ratio, 155 for slope and 154 for slope-ratio models were extracted and BPNN and non-parametric statistical classifier were used for classification. The classification was also assessed with the top 20 features.

The top 20 features of the slope-ratio model were more robust and had maximum classification accuracy than the ratio and slope models in classifying the cereals, oilseeds, pulses and specialty seeds. The effects of growing region and crop-year, moisture content and foreign material content were not classified accurately. The BPNN classifier classified more accurately than the statistical classifier for different moisture contents of wheat. But the statistical classifier was more successful in classifying the foreign material content than the BPNN classifier. The features, which contributed to classification, were spread throughout the spectrum making it difficult to use filters for practical applications. The majority of the features were, however, from the visible and near infrared range. The reflectance data have a promise for application for quick classification of bulk grain at grain elevators.

REFERENCES

Anonymous. 2001. Diffuse Reflectance Accessory for the Cary 4, 400, 5 and 500. Varian Canada Inc. Mill creek. Mississauga. Ontario

ASAE. 1997. Moisture measurement – unground grain and seeds. ASAE S352.2. In ASAE Standards 1997, 555. St. Joseph, MI: ASAE.

Bilanski, W. K., C. L. Pen, and D. R. Fuzzen. 1984. Apple bruise detection using optical reflectance parameters. Canadian Agricultural Engineering 26(2): 111-114

Bittner, D. R. and K. H. Norris. 1968. Optical properties of selected fruits vs. maturity. Transactions of the ASAE 11(4): 534-536

Canada Grains Council. 1982. Grain Grading for Efficiency and Profit. Canada Grains Council, Winnipeg, MB.

Casady, W. W., M.. R . Paulsen, and J. B. Sinclair. 1993. Optical properties of damaged soybean seeds. Transactions of the ASAE 36(3): 943-947.

Chen, C., Y. P. Chiang, and Y. Pomeranz. 1989. Image analysis and characterization of cereal grains with a laser range finder and camera contour extractor. Cereal Chemistry 66(6): 466-470.

CIGI. 1993. Grains and Oilseeds- Handling, Marketing, Processing, 4th ed. Winnipeg, MB. Canadian International Grains Institute.

Delwiche, S. R. 1995. Single wheat kernel analysis by near-infra red transmittance: protein content. Cereal Chemistry 72(1): 11-16.

Delwiche, S. R. and K. H. Norris. 1993. Classification of hard red wheat by near-infra red diffuse reflectance spectroscopy. Cereal Chemistry 70(1): 29-35.

Delwiche, S. R., Y. R. Chen, and W. R. Hruschka. 1995b. Differentiation of hard red wheat by near-infra red analysis of bulk samples. Cereal Chemistry. 72(3): 243-247.

Dowell, F. E. 1992. Identifying undamaged and damaged peanut kernels using tristimulus values and spectral reflectance. Transactions of the ASAE 35(3): 931-937.

Dowell, F. E., J. E. Throne, and J. E. Baker. 1998. Automated nondestructive detection of internal insects infestations of wheat kernels using near-infra red reflectance spectroscopy. Journal of Economic Entomology 91(4): 899-904.

Eu, M. T. 1997. Reflectance characteristics of bulk grains using a spectrophotometer. Unpublished M.Sc Thesis. Department of Biosystems Engineering, University of Manitoba, Winnipeg, MB.

Hawk, A. L., H. H. Kaufmann, and C. A. Watson. 1970. Reflectance characteristics of various grains. Cereal Science Today 15(11): 381-384.

Heinemann, P. H., R. Hughes, C. T. Morrow, H. J. Sommer III, R. B. Beelman, and P. J. Wuest. 1994. Grading of mushrooms using a machine vision system. Transactions of the ASAE 37(5): 1671-1677.

Howarth, M. S., S. W. Searcy, and G. S. Birth. 1990. Reflectance characteristics of fresh-market carrots. Transactions of the ASAE 33(3): 961-964.

Jayas, D.S., J.Paliwal, and N.S. Visen. 2000. Multi-Layer neural networks for image analysis of agricultural products. Journal of Agricultural Engineering Research 77(2):119-128

Kohler,J.M. 1991. Study of commercial grades of Canada red Spring wheat by digital image processing. Unpublished M.Sc Thesis. Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, MB.

Karunakaran, C. 2002. Soft X-Ray inspection of wheat kernels to detect infestations by stored grain insects. Unpublished Ph.D. Thesis. Department of Biosystems Engineering, University of Manitoba, Winnipeg, MB.

Luo, X. 1997. Color image analysis for cereal grain classification. Unpublished Ph.D. Thesis, Department of Biosystems Engineering, University of Manitoba, Winnipeg, MB.

Majumdar, S., D. S. Jayas, S. J. Symons, and N. R. Bulley. 1996a. Textural features for automated grain identification. CSAE Paper No. 96-602. Saskatoon, SK: CSAE.

Majumdar, S., D. S. Jayas, J. L.Hehn, and N. R. Bulley. 1996b. Classification of various grains using optical properties. Canadian Agricultural Engineering. 38(2): 139-144.

Majumdar, S., and D. S. Jayas, 1999. Classification of bulk samples of cereal grains using machine vision. Journal of Agricultural Engineering Research 73:35-47

Neuman, M., H. D. Sapirstein, E. Shwedyk, and W. Bushuk. 1986. Discrimination of wheat class and variety by digital image analysis of whole grain samples. Journal of Cereal Science 6: 125-132.

Paliwal,J. 2002. Digital image analysis of grain samples for potential use in grain cleaning. Unpublished Ph.D. Thesis. Department of Biosystems Engineering, University of Manitoba, Winnipeg, MB.

Paliwal,J., N.S. Visen and D.S. Jayas. 2001. Evaluation of neural network architectures for cereal grain classification using morphological features. Journal of Agricultural Engineering Research 79(4):361-370.

Panford, J. A. 1987. Application of near-infra red reflectance spectroscopy in North America. Chapter 10. In Near-Infrared Technology in the Agricultural and Food Industries. Williams, P. and K. Norris (ed.) American Association of Cereal Chemists, Inc., St. Paul, Minnesota, U.S.A.

Putnam, D.F. and R.G. Putnam. 1970. *Canada : a regional analysis*. Toronto, ON: J.M. Dent and Sons, Inc.

Reyer. Z., Q .Yang., E. Gracia-Pardo, and R Bull. 1995. Use of spectral information and machine vision for bruise detection on peaches and apricots. Journal of Agricultural Engineering Research, 1996(63): 323-332.

SAS.1990. SAS User's Guide: Statistics. Raleigh. NC: Statistical Analysis System, Inc.

Shatadal, P., D. S. Jayas, and N. R. Bulley. 1994. Digital image analysis for software separation and classification of touching grains: II. Classification. Transactions of the ASAE 38(2): 645-649.

Shatadal, P., D. S. Jayas, J. L. Hehn, and N. R. Bulley. 1995. Seed classification using machine vision. Canadian Agricultural Engineering 37(3): 163-167.

Singh, N. and M. J. Delwiche. 1994. Machine vision methods for defect sorting stone fruit. Transactions of the ASAE 37(6): 1989-1997.

Stermer, R. A., Y. Pomeranz, and R. J. Mc Ginty. 1977. Infrared reflectance spectroscopy for estimation of moisture of whole grain. Cereal Chemistry 54(2): 345-351.

Symons, S. J. and R. G. Fulcher. 1988a. Determination of wheat kernel morphological variation by digital image analysis: I. Variation in Eastern Canadian milling quality wheats. Journal of Cereal Science 8: 211-218.

Symons, S. J. and R. G. Fulcher. 1988b. Determination of wheat kernel morphological variation by digital image analysis: II. Variation in cultivars of soft white winter wheats. Journal of Cereal Science 8: 219-229.

Tao, Y., P. H. Heinemann, Z. Varghese, C. T. Morrow, and H. J. Sommer III. 1995. Machine vision for color inspection of potatoes and apples. Transactions of the ASAE 38(5): 1555-1561.

Thomson, W. H. and Y. Pomeranz. 1991. Classification of wheat kernels using three dimensional image analysis. Cereal Chemistry 68(4): 357-361.

Upchurch, B.L., H. A. Affeldt, W. R. Hruschka, K. H. Norris, and J. A. Throop. 1990. Spectrophotometric study of bruises on whole, 'Red Delicious' apples. Transactions of the ASAE 33(2): 585-589.

Williams, P. C. 1975. Application of near-infrared reflectance spectroscopy to analysis of cereal grains and oilseeds. Cereal Chemistry 52: 561-576.

Williams. P. and K. Norris. 1987. *Near-Infrared Technology in the Agricultural and Food Industries.* American Association of Cereal Chemists, Inc., St. Paul, Minnesota, U.S.A.

Visen,N.S. 2002. Machine vision based grain handling system. Unpublished Ph.D. Thesis. Department of Biosystems Engineering, University of Manitoba, Winnipeg, MB

APPENDIX A: A Sample of the Reflectance Curve for Each Sample.

The legend of each figure is listed in corresponding order based on the end of each curve.

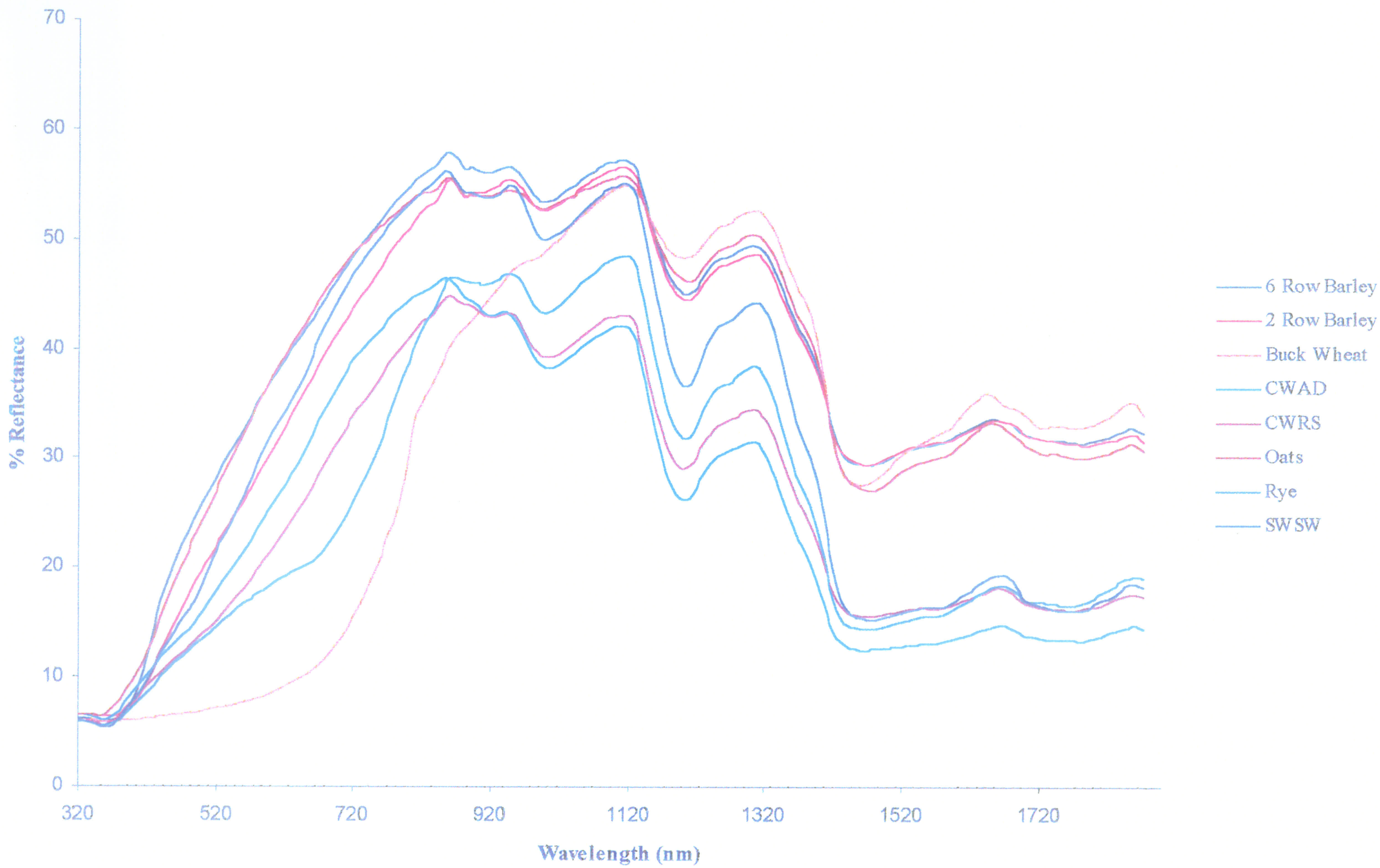

Fig A1. Reflectance characteristics of cereals and buckwheat based on one sample of each cereal and buckwheat

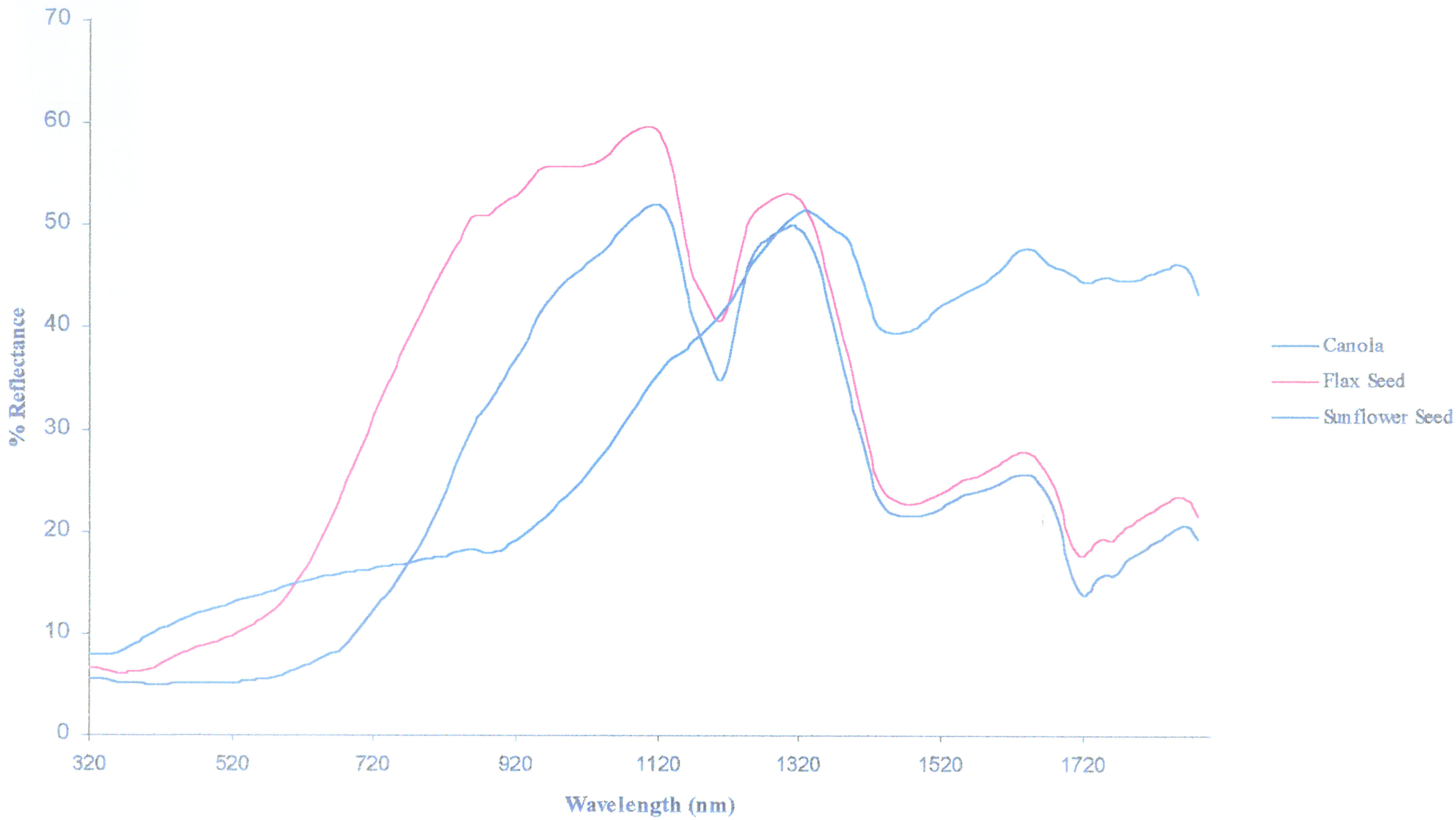

Fig A2. Reflectance characteristics of oilseeds based on one sample of each oilseed

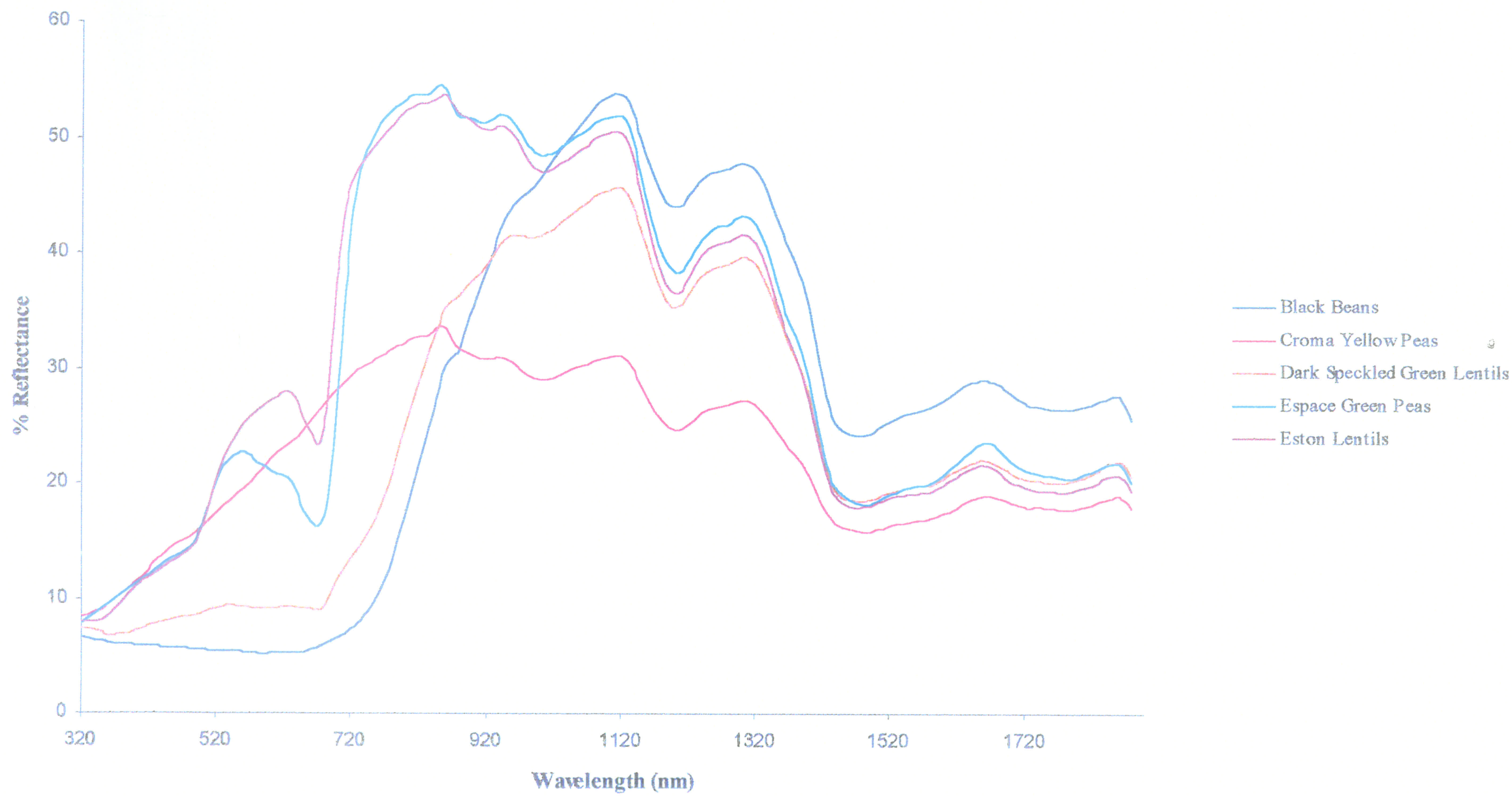

Fig A3a. Reflectance characteristics of pulses based on one sample of each pulse (five pulses shown on this graph and the remaining five shown in Fig A3b)

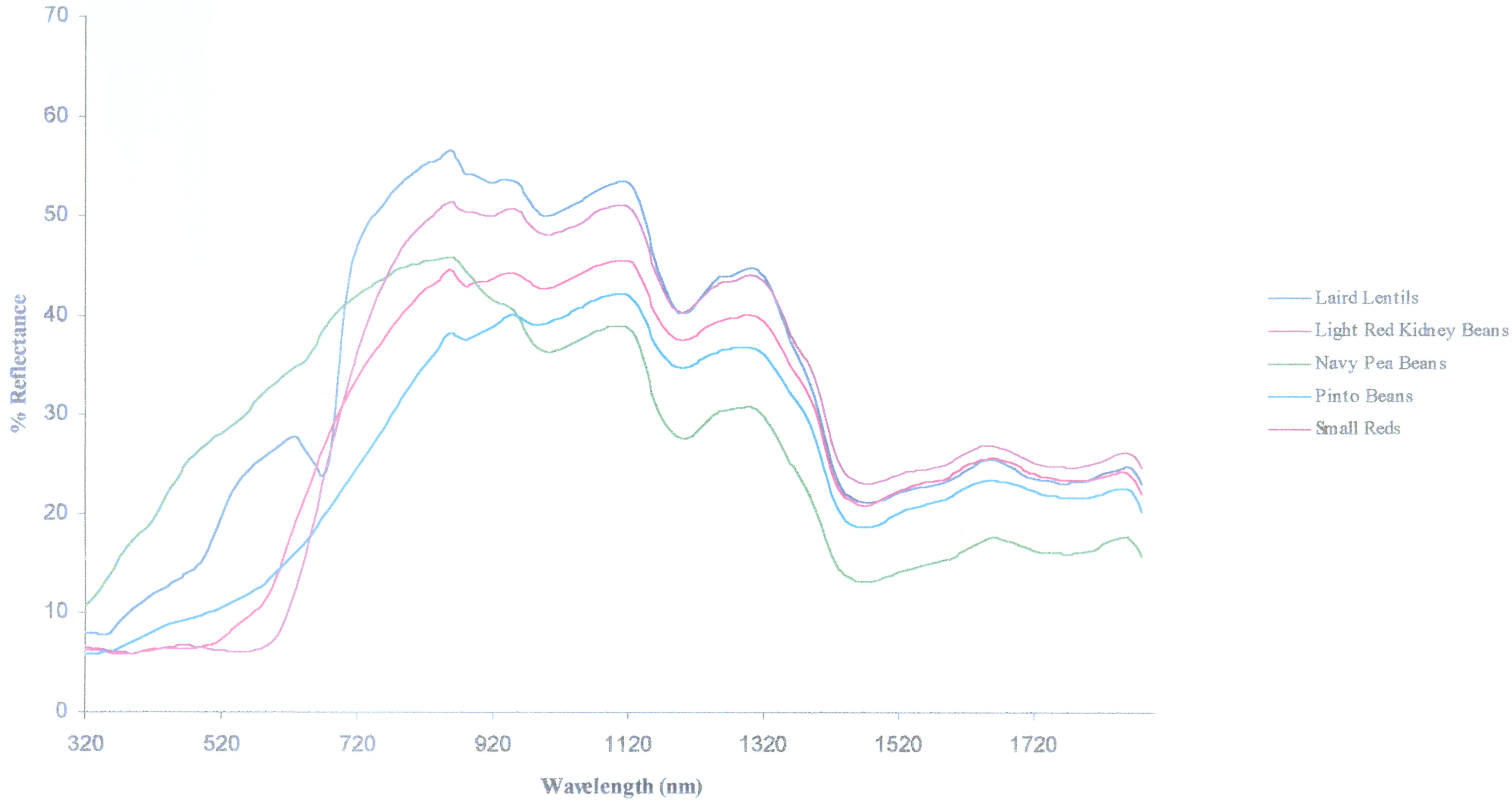

Fig A3b. Reflectance characteristics of pulses based on one sample of each pulse (five pulses shown on this graph and the remaining five shown in Fig A3a)

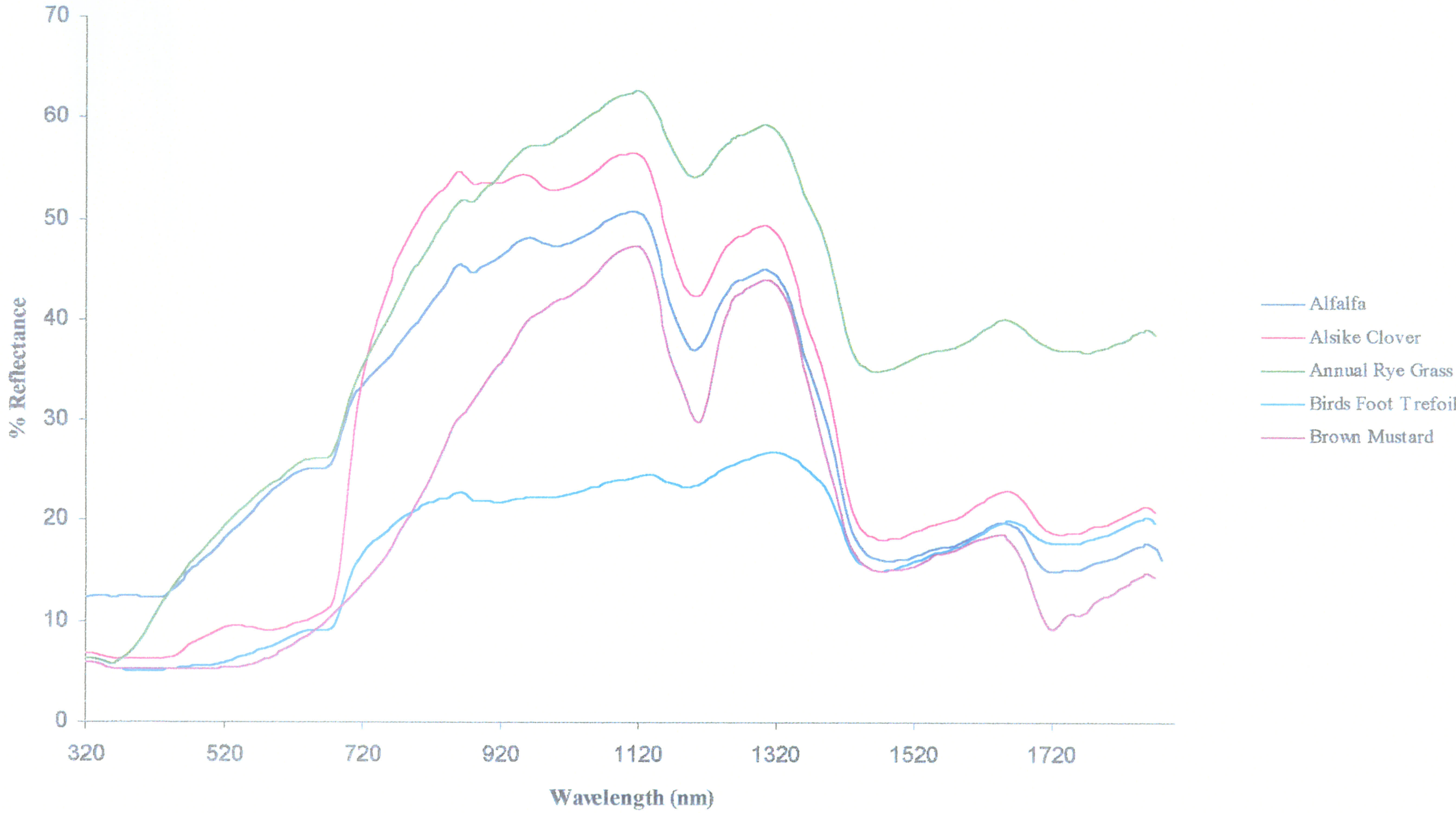

Fig A4a. Reflectance characteristics of specialty seeds based on one sample of each specialty seeds (five specialty seeds shown in each of the Fig A4a, A4b, A4c, A4d, and A4e)

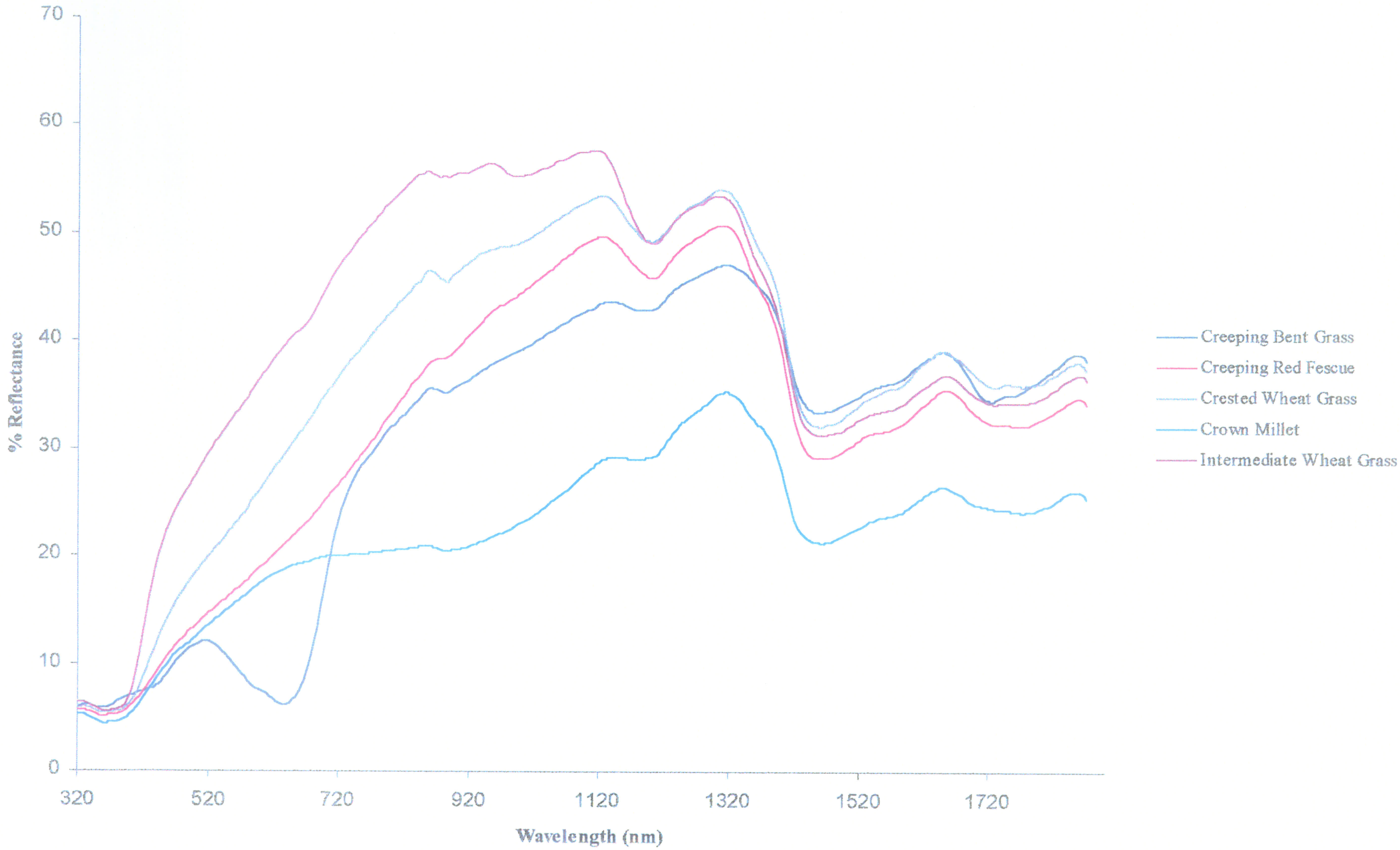

Fig A4b. Reflectance characteristics of specialty seeds based on one sample of each specialty seeds (five specialty seeds shown in each of the Fig A4a, A4b, A4c, A4d, and A4e)

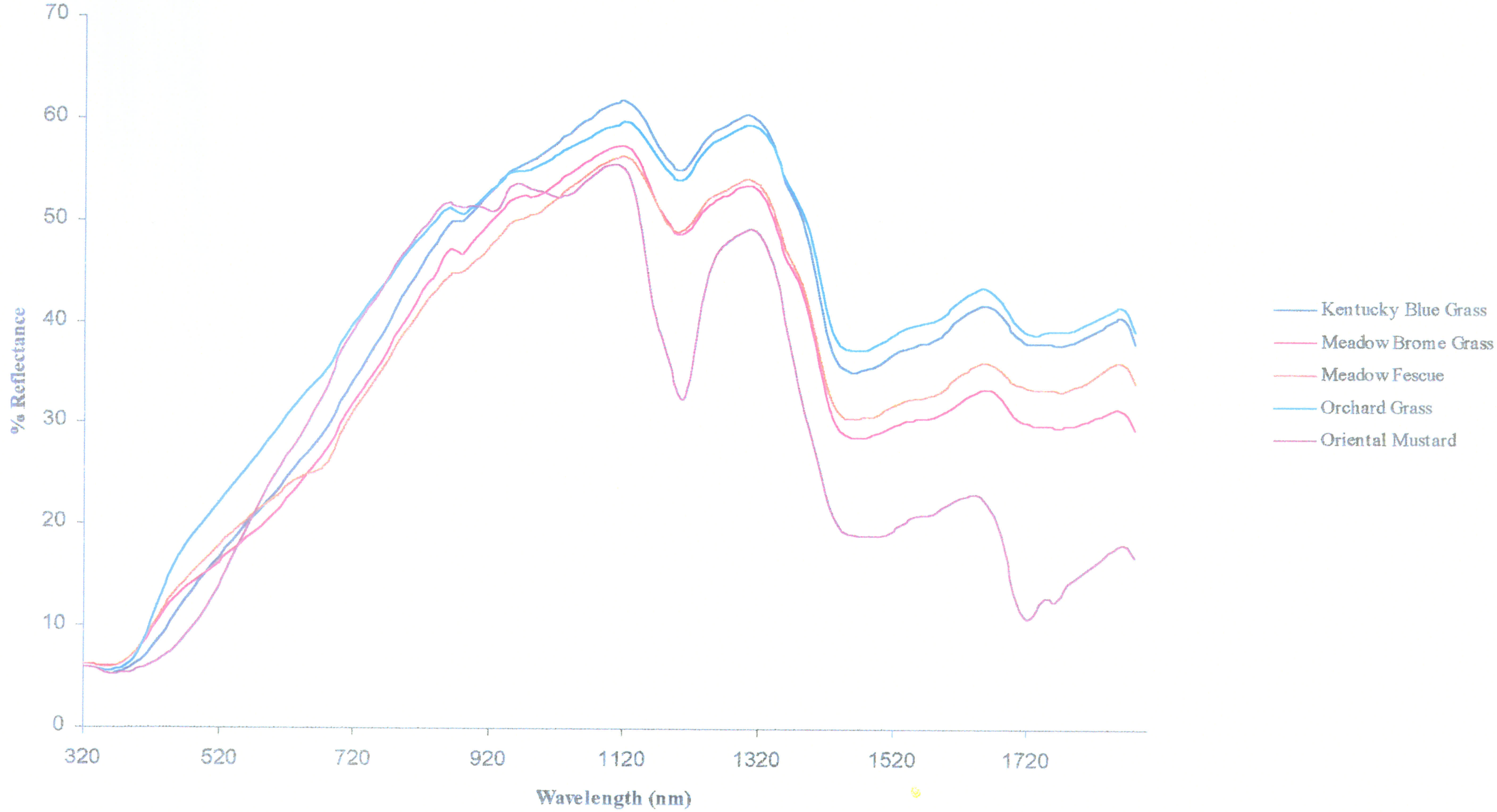

Fig A4c. Reflectance characteristics of specialty seeds based on one sample of each specialty seeds (five specialty seeds shown in each of the Fig A4a, A4b, A4c, A4d, and A4e)

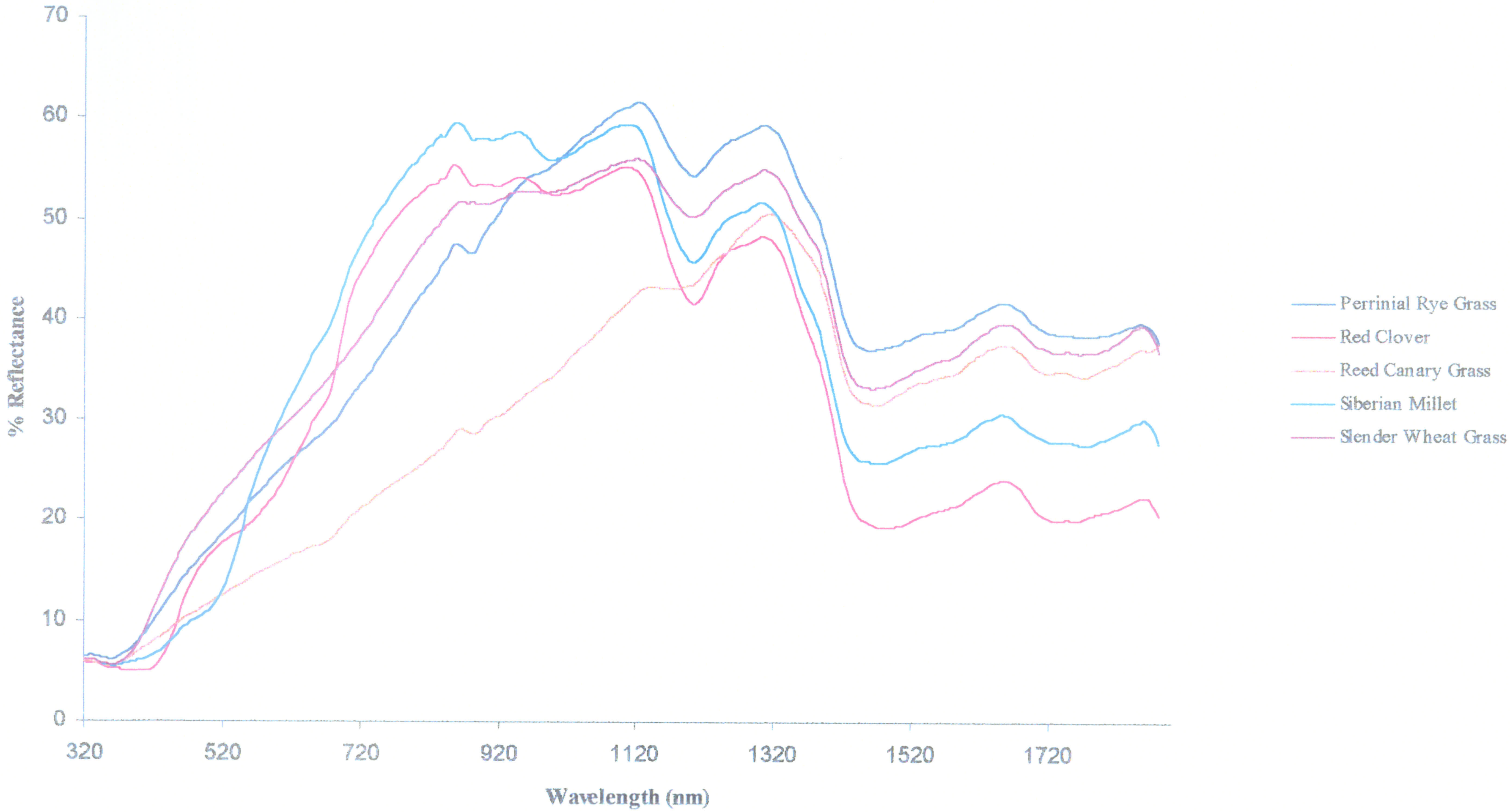

Fig A4d. Reflectance characteristics of specialty seeds based on one sample of each specialty seeds (five specialty seeds shown in each of the Fig A4a, A4b, A4c, A4d, and A4e)

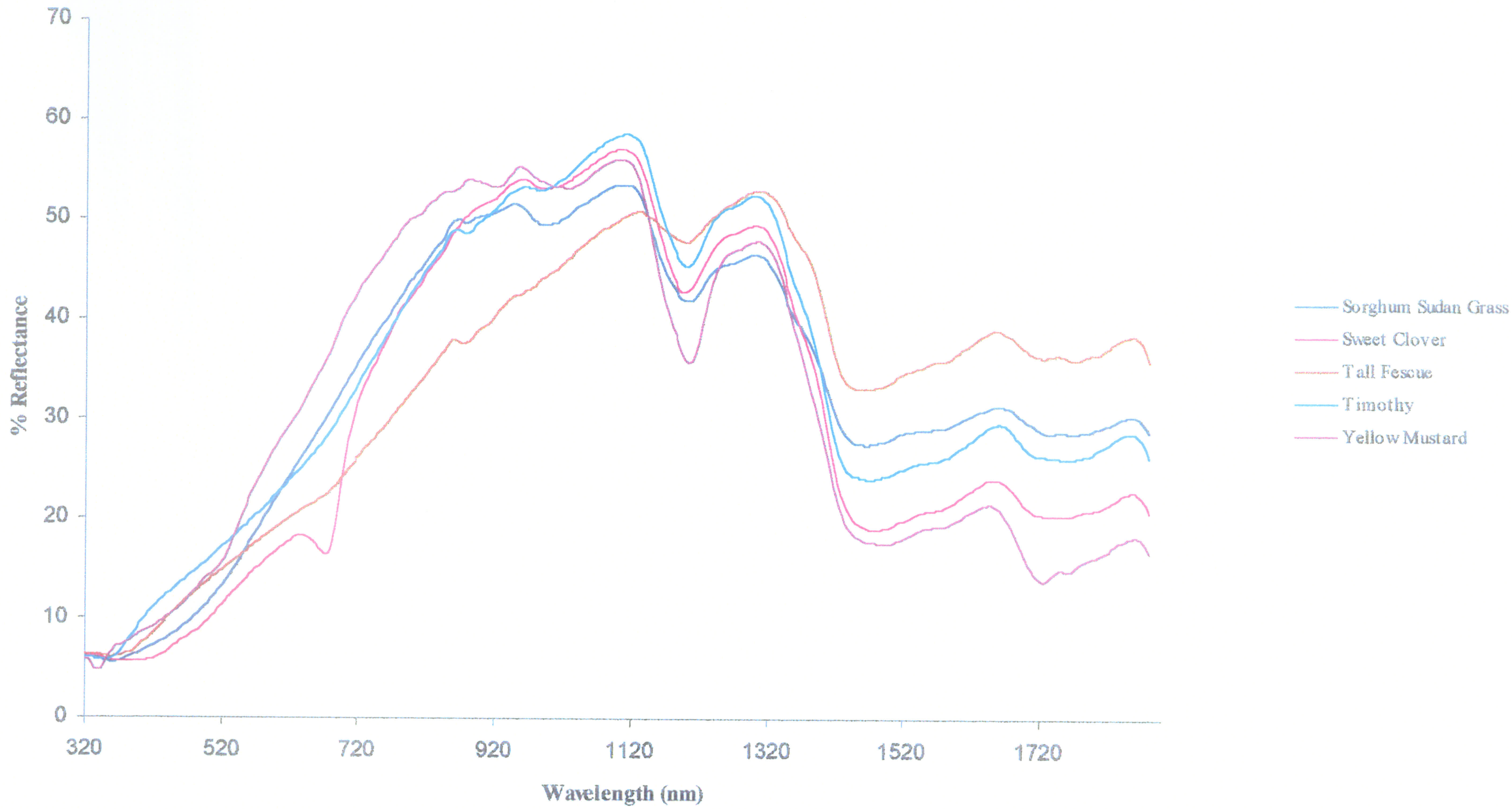

Fig A4e. Reflectance characteristics of specialty seeds based on one sample of each specialty seeds (five specialty seeds shown in each of the Fig A4a, A4b, A4c, A4d, and A4e)

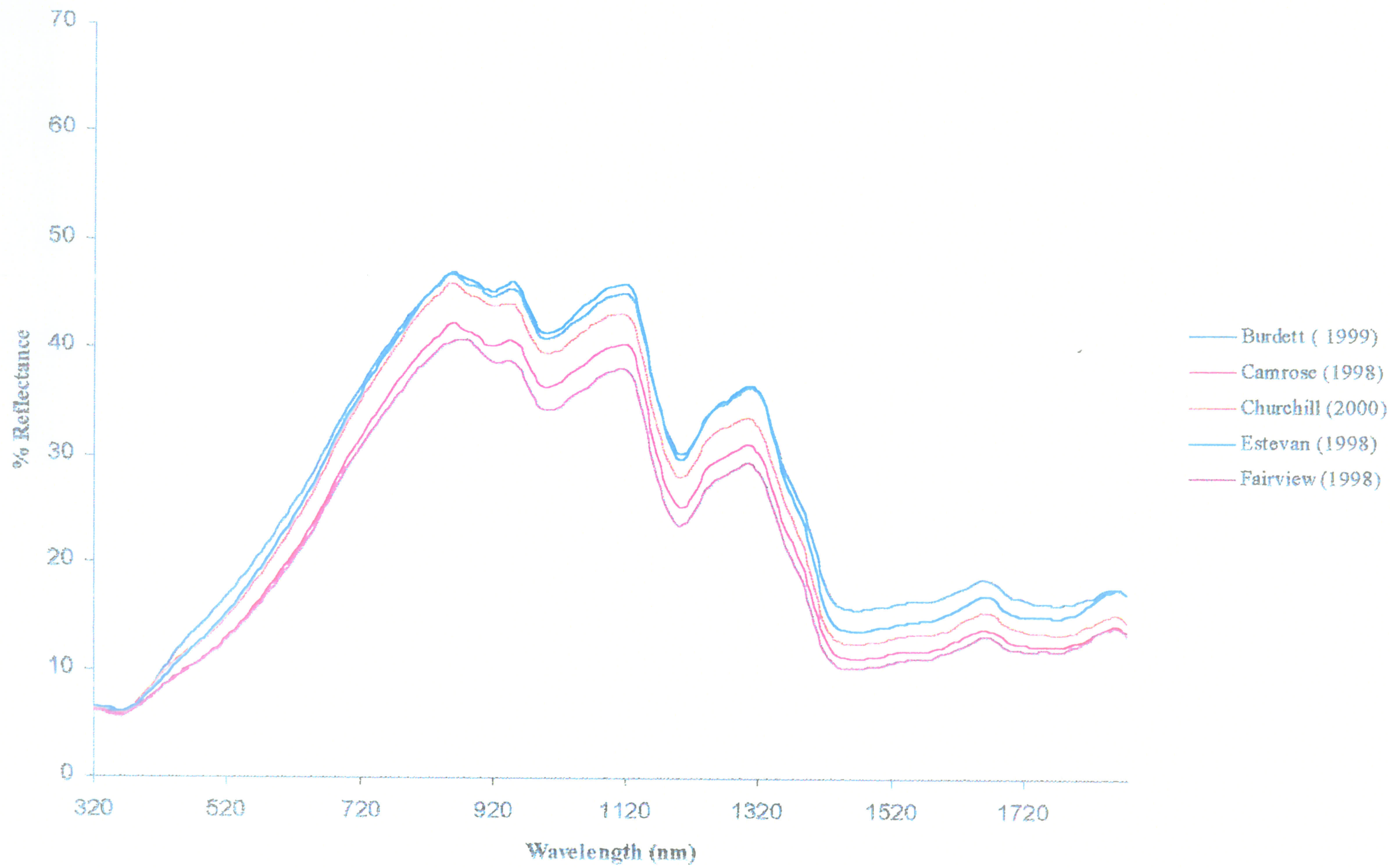

Fig A5a.Reflectance characteristics of CWRS wheat based on one sample of CWRS wheat from indicated growing regions and years (five growing regions and years are shown on this graph and in Fig A5b, and A5c)

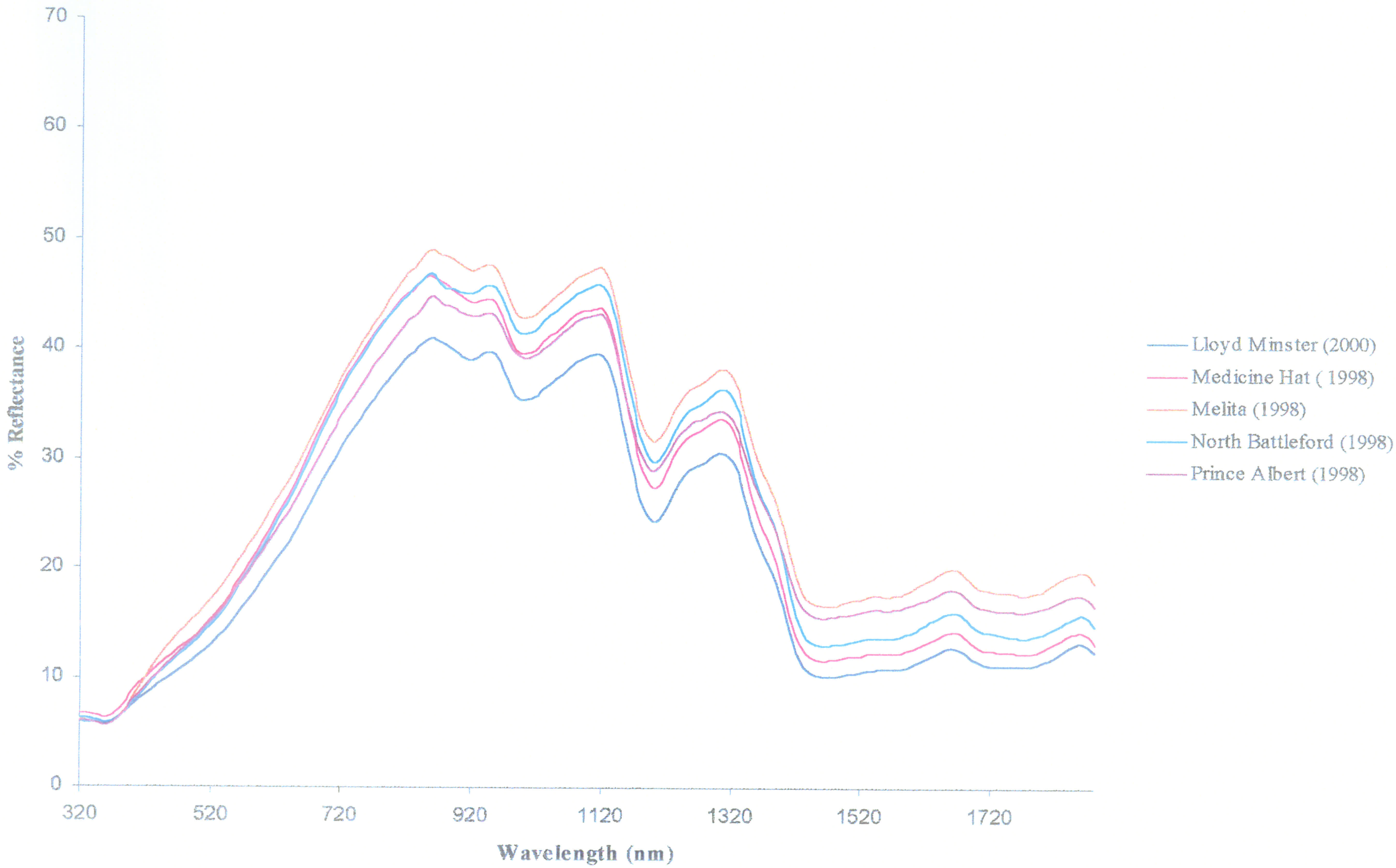

Fig A5b.Reflectance characteristics of CWRS wheat based on one sample of CWRS wheat from indicated growing regions and years (five growing regions and years are shown on this graph and in Fig A5a, and A5c)

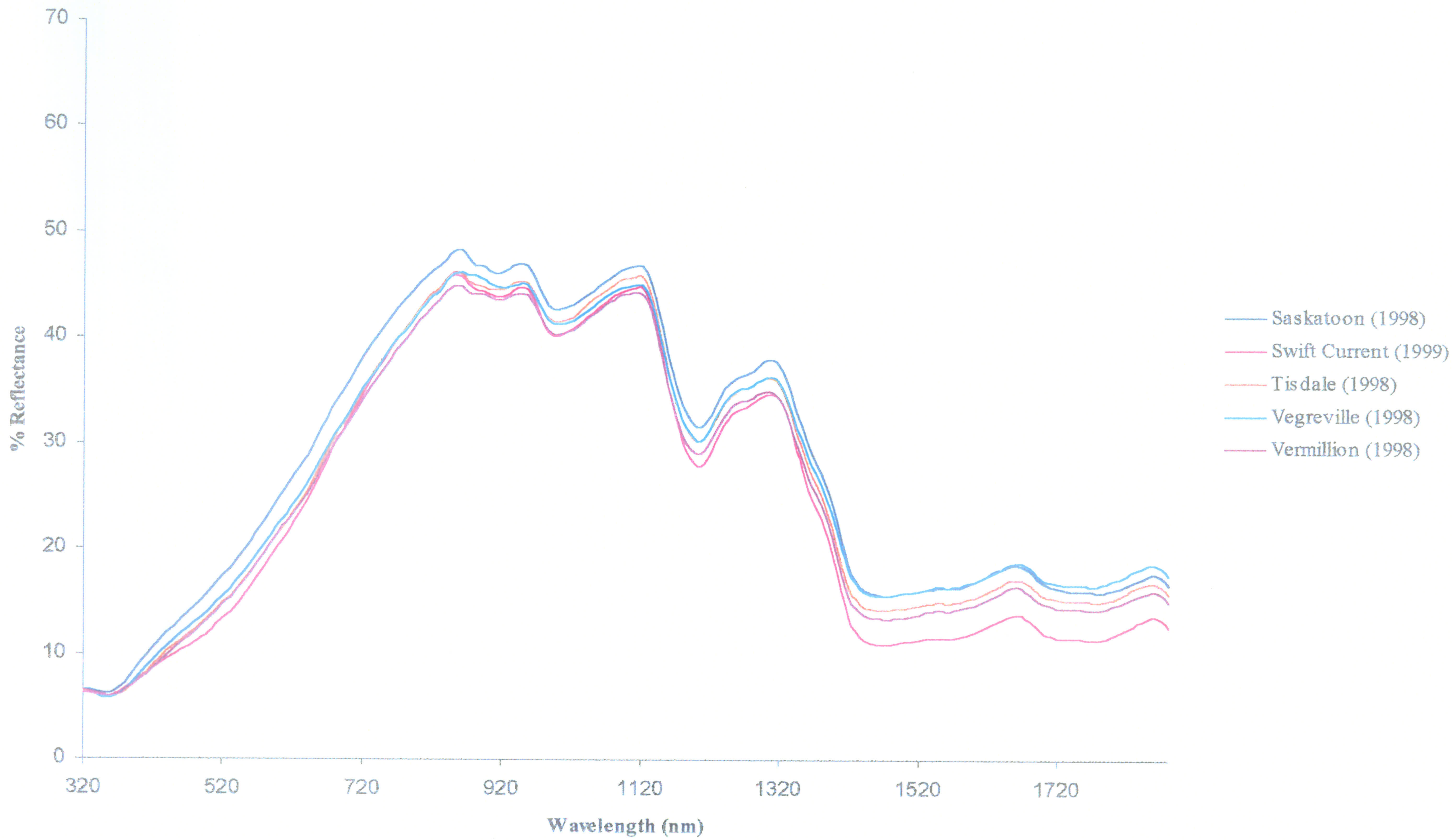

Fig A5c.Reflectance characteristics of CWRS wheat based on one sample of CWRS wheat from indicated growing regions and years (five growing regions and years are shown on this graph and in Fig A5a, and A5b)

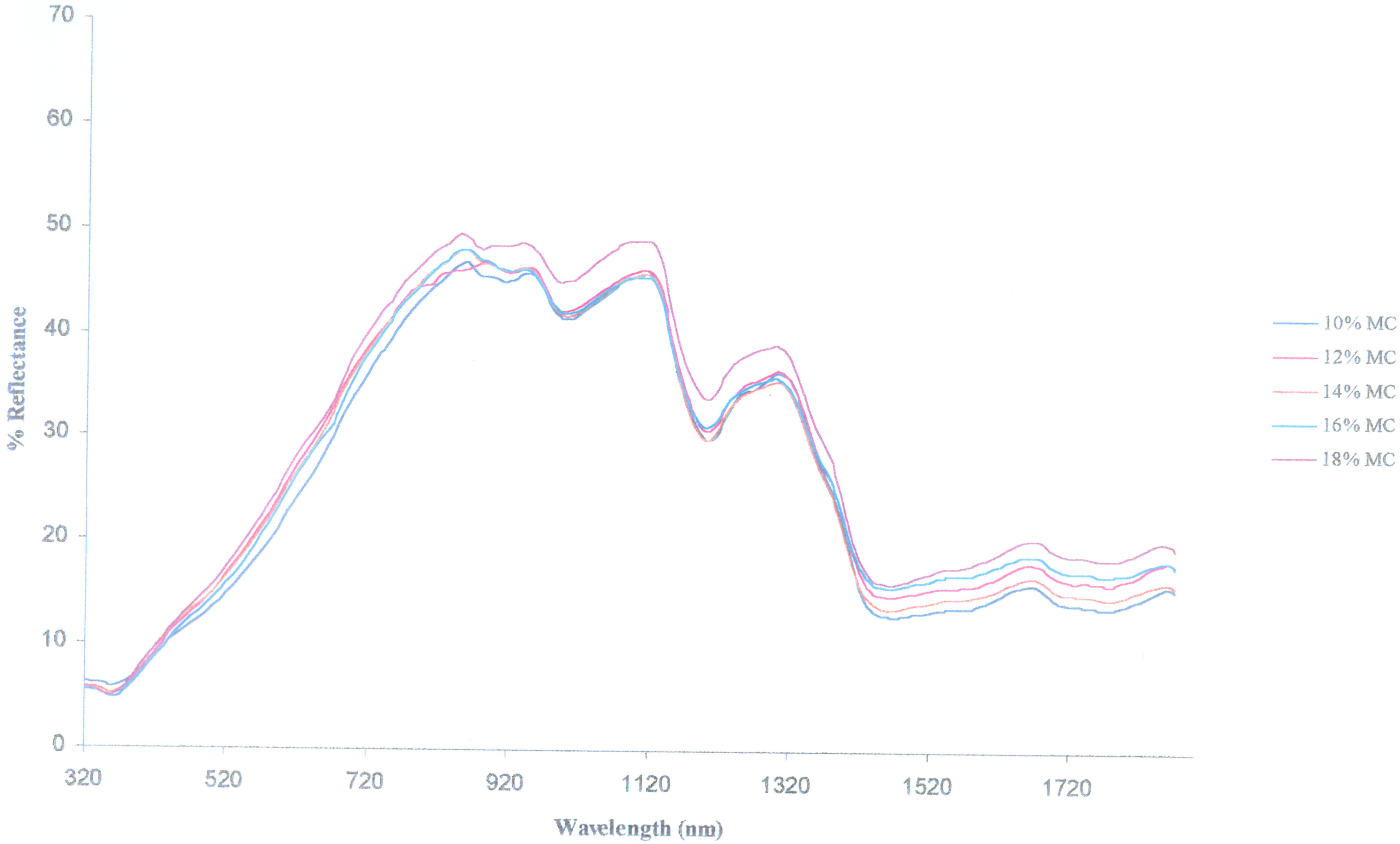

Fig A6.Reflectance characteristics of CWRS wheat based on one sample for each moisture content

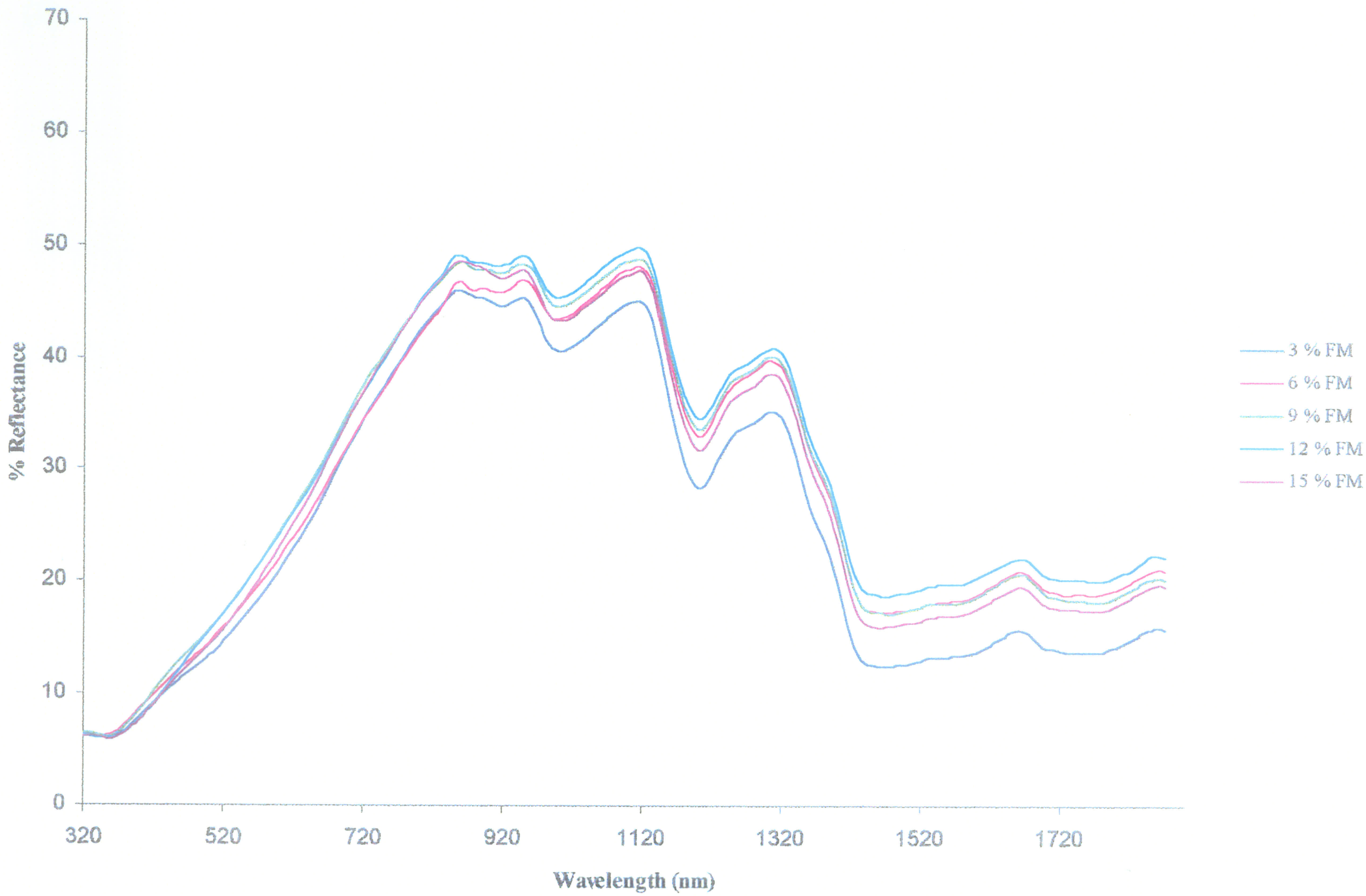

Fig A7.Reflectance characteristics of CWRS wheat based on one sample for each foreign material content

APPENDIX B: Confusion Matrices.

List of seeds for Tables B1a and B1b

1 - 2 row barley
2 - 6 row barley
3 - CWRS wheat
4 - CWAD wheat
5 - Buckwheat
6 - Oats
7 - Rye
8 - SWSW
9 - Canola
10 - Flax
11 - Sunflower
12 - Black beans
13 - Croma yellow peas
14 - Dark speckled green lentils
15 - Espace green peas
16 - Eston lentils
17 - Laird lentils
18 - Light red kidney beans
19 - Navy pea beans
20 - Pinto beans
21 - Small red kidney beans

Table B1a. Confusion matrix of the twenty feature slope-model with 21 classes for the hold-out method (Non-parametric estimation) for cereals and buckwheat, pulses and oilseeds

Class (to) → (from) ↓	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	1	2	0	0	0	15	0	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	18	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	18	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	18	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	18	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	18	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0	16	1	1	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	17	0	0	1
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18	0	0
20	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	16	0
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18

Table B1b. Confusion matrix of the twenty feature slope model with 21 classes for the four layer BPNN for cereals and buckwheat, pulses and oilseeds

Class (to) → (from) ↓	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	17	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	18	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	18	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	18	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	18	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	16	0	0	0	0	1	1	0	0
14	0	0	0	0	0	0	0	0	0	0	0	1	0	15	1	0	0	0	0	1	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	17	0	0	1
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18	0	0
20	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	16	0
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	17

List of locations and crop- years for the CWRS wheat samples in Tables B2a, B2b, B2c, B2d, B2e, and B2f

1 - Burdett (1999)
2 - Camrose (1998)
3 - Churchill (2000)
4 - Estevan (1998)
5 - Fairview (1998)
6 - Lloyd Minster (2000)
7 - Medicine Hat (1998)
8 - Melita (1998)
9 – North Battleford (1998)
10 – Prince Albert (1998)
11 – Saskatoon (1998)
12 – Swift Current (1999)
13 – Tisdale (1998)
14 – Vegerville (1998)
15 – Vermillion (1998)
16 – Not classified

Table B2a. Confusion matrix of the twenty feature ratio-model with 25 classes for hold-out method (Non-parametric estimation) of regions and years

Class (to) → (from) ↓	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	3	1	0	0	0	1	0	0	2	0	1	2	0	0	0	8
2	1	0	0	3	0	0	2	0	0	0	0	0	0	0	0	12
3	0	1	2	0	1	0	0	2	0	0	0	1	0	0	1	10
4	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	15
5	0	0	0	0	3	0	0	0	0	0	1	0	0	3	0	11
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18
7	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	17
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18
9	1	2	0	1	0	0	0	0	7	0	0	0	0	1	0	6
10	1	1	0	1	0	0	0	0	0	3	1	3	0	0	4	4
11	4	3	0	1	0	0	0	0	0	0	6	0	0	0	0	4
12	0	5	0	2	0	0	0	0	3	1	2	1	0	0	3	1
13	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	16
14	3	2	0	0	0	0	1	0	0	0	0	1	0	3	0	8
15	5	3	0	0	0	4	0	0	0	0	0	0	1	0	5	0

Table B2b. Confusion matrix of the twenty-feature ratio-model with 25 classes for the four layer BPNN of regions and years

Class (to) → (from) ↓	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	2	1	0	0	1	1	0	0	1	0	3	0	1	0	1	7
2	1	1	0	1	0	1	1	0	1	0	3	0	0	2	0	7
3	1	0	2	0	1	0	0	0	1	0	0	0	1	0	1	11
4	1	1	0	3	0	0	0	0	4	0	0	0	0	0	0	9
5	0	1	0	0	3	0	0	0	2	0	1	0	0	3	0	8
6	0	1	0	0	0	2	0	0	0	0	0	0	0	0	0	15
7	1	2	0	0	0	0	1	0	2	0	3	0	0	0	0	9
8	0	4	1	0	0	0	0	0	4	0	0	0	0	0	0	9
9	2	1	0	1	0	0	0	0	8	0	0	0	0	2	0	4
10	1	1	0	0	0	1	0	0	0	1	3	3	0	0	4	4
11	1	1	0	1	0	0	0	0	2	0	2	0	0	0	0	11
12	0	1	0	1	0	0	0	0	1	1	2	1	0	0	3	8
13	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	15
14	1	2	0	0	0	2	1	0	2	0	0	1	0	3	0	6
15	1	0	0	0	0	2	0	0	2	0	3	2	0	0	2	6

Table B2c. Confusion matrix of the twenty-feature slope-model with 25 classes for hold-out method (Non-parametric estimation) of regions and years

Class (to) → (from) ↓	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	3	1	0	1	0	1	0	0	2	0	1	0	1	1	0	7
2	1	4	3	3	1	0	1	0	0	1	2	1	0	1	0	0
3	2	1	2	0	1	0	1	0	3	0	3	0	0	3	1	1
4	4	0	0	5	0	0	0	0	0	1	2	0	0	1	0	5
5	4	0	0	0	2	0	1	0	0	1	4	0	0	0	0	6
6	3	0	0	0	0	3	0	0	0	3	0	4	3	0	0	2
7	4	0	0	0	0	0	1	0	0	0	0	0	0	1	0	12
8	2	0	1	0	0	0	0	2	0	3	0	0	0	0	0	10
9	2	2	0	1	0	0	0	0	10	0	0	3	0	0	0	0
10	3	1	0	1	0	0	0	0	0	6	1	0	0	1	4	1
11	4	3	0	1	0	0	0	0	0	0	6	0	0	4	0	0
12	4	1	0	1	0	0	0	0	1	1	3	3	0	1	3	0
13	1	0	0	0	0	0	0	0	0	0	1	0	3	0	0	13
14	4	2	0	0	0	0	1	0	0	3	0	0	0	4	1	3
15	1	1	0	0	0	1	0	0	0	1	3	0	1	0	6	4

Table B2d. Confusion matrix of the twenty-feature slope-model with 25 classes for the four layer BPNN of regions and years

Class (to) → (from) ↓	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	1	2	1	0	1	1	0	1	1	0	1	2	1	2	0	4
2	3	5	1	1	0	0	1	1	1	0	1	1	0	0	3	0
3	1	0	11	0	1	0	0	0	3	0	0	1	0	0	1	0
4	3	0	0	5	0	1	0	0	0	0	1	0	1	0	3	4
5	0	0	0	0	0	0	0	0	0	0	1	0	0	3	0	14
6	4	0	0	0	0	4	0	0	0	0	0	0	0	0	4	6
7	4	0	0	0	0	3	3	0	4	0	0	1	0	0	3	0
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18
9	3	1	0	1	0	0	0	1	10	1	0	0	0	1	0	0
10	1	1	0	1	0	0	0	0	0	4	1	1	0	0	1	8
11	2	2	0	1	0	0	0	0	3	0	4	0	1	0	4	1
12	5	2	0	0	0	0	0	0	1	0	1	5	0	0	1	3
13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18

Table B2e. Confusion matrix of the twenty-feature slope ratio-model with 25 classes for hold-out method (Non-parametric estimation) of regions and years

Class (to) → (from) ↓	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	4	1	0	0	0	1	0	0	1	0	1	2	0	3	3	2
2	1	4	0	3	0	2	1	0	0	0	0	0	0	4	3	0
3	2	1	10	0	2	0	0	0	1	0	0	0	0	1	1	0
4	4	0	1	0	0	0	1	0	2	0	0	0	0	0	6	4
5	1	1	0	0	4	0	0	0	1	0	1	0	0	3	0	7
6	2	0	0	0	0	9	0	0	1	0	0	0	0	3	2	1
7	2	0	1	0	0	0	6	0	1	0	0	3	0	0	5	0
8	2	0	0	2	0	0	0	3	0	0	0	0	0	3	0	8
9	2	1	1	0	0	0	0	0	14	0	0	0	0	0	0	0
10	1	1	1	1	0	1	0	0	0	5	1	0	0	3	4	0
11	3	3	0	1	0	0	0	0	2	0	4	1	0	4	0	0
12	0	2	1	2	0	1	0	0	1	1	2	5	0	0	3	0
13	0	3	0	0	0	0	2	2	1	0	0	0	5	0	3	2
14	3	2	0	0	0	0	1	0	1	0	0	0	0	8	2	1
15	1	2	0	1	1	2	0	0	1	0	0	0	1	0	9	0

Table B2f. Confusion matrix of the twenty-feature slope ratio-model with 25 classes for the four layer BPNN of regions and years

Class (to) → (from) ↓	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	6	3	0	0	0	0	0	0	6	0	0	0	0	0	3	0
2	3	9	0	1	0	0	0	0	3	0	0	1	0	0	0	1
3	2	0	8	0	1	0	0	0	3	0	0	1	0	0	1	2
4	3	0	0	5	0	0	0	0	6	0	0	0	0	0	2	2
5	3	0	0	0	6	0	0	0	3	0	0	0	0	3	3	0
6	6	0	0	0	0	7	0	0	2	0	0	0	0	0	0	3
7	6	0	0	0	0	0	4	0	3	0	0	0	0	0	3	2
8	3	0	0	0	0	0	0	4	0	0	0	0	0	0	0	11
9	1	2	0	1	0	0	0	0	13	0	0	0	0	0	1	0
10	3	1	3	1	0	0	0	0	3	3	1	1	0	0	1	1
11	1	0	0	0	0	0	0	0	0	0	7	0	0	0	0	10
12	4	0	0	0	0	0	0	0	3	0	2	5	0	0	3	1
13	3	0	0	0	0	0	0	0	0	0	0	0	4	0	6	5
14	4	2	0	0	0	0	1	0	0	0	0	0	0	6	0	5
15	3	0	0	0	0	0	0	0	4	0	0	0	1	0	9	1

Table B3a. Confusion matrix of the twenty-feature ratio-model with 25 classes for hold-out method (Non-parametric estimation) for moisture content of CWRS wheat

Class (to) → (from) ↓	MC 10	MC 12	MC 14	MC 16	MC 18
MC 10	18	0	0	0	0
MC 12	2	15	1	0	0
MC 14	0	0	16	2	0
MC 16	1	0	1	16	0
MC 18	0	0	0	0	18

Table B3b. Confusion matrix of the twenty-feature slope ratio-model with 25 classes for the four layer BPNN for moisture content of CWRS wheat

Class (to) → (from) ↓	MC 10	MC 12	MC 14	MC 16	MC 18
MC 10	18	0	0	0	0
MC 12	1	15	2	0	0
MC 14	0	1	16	1	0
MC 16	0	0	1	17	0
MC 18	0	1	0	0	17

Table B3c. Confusion matrix of the twenty feature slope-model with 25 classes for hold-out method (Non-parametric estimation) for moisture content of CWRS wheat

Class (to) → (from) ↓	MC 10	MC 12	MC 14	MC 16	MC 18
MC 10	14	3	1	0	0
MC 12	6	9	2	1	0
MC 14	0	3	11	4	0
MC 16	0	0	1	17	0
MC 18	0	0	0	1	17

Table B3d. Confusion matrix of the twenty-feature slope model with 25 classes for the four layer BPNN for moisture content of CWRS wheat

Class (to) → (from) ↓	MC 10	MC 12	MC 14	MC 16	MC 18
MC 10	14	1	3	0	0
MC 12	4	9	3	2	0
MC 14	0	2	8	5	3
MC 16	2	2	3	9	2
MC 18	1	0	0	3	14

Table B3e. Confusion matrix of the twenty-feature slope ratio-model with 25 classes for hold-out method (Non-parametric estimation) for moisture content of CWRS wheat

Class (to) → (from) ↓	MC 10	MC 12	MC 14	MC 16	MC 18
MC 10	17	`1	0	0	0
MC 12	0	14	2	0	0
MC 14	1	2	15	0	0
MC 16	0	1	3	14	0
MC 18	0	1	1	0	16

Table B3f. Confusion matrix of the twenty-feature slope ratio-model with 25 classes for the four layer BPNN for moisture content of CWRS wheat

Class (to) → (from) ↓	MC 10	MC 12	MC 14	MC 16	MC 18
MC 10	17	1	0	0	0
MC 12	2	15	1	0	0
MC 14	0	2	15	1	0
MC 16	0	1	2	15	0
MC 18	0	0	0	0	18

Table B4a. Confusion matrix of the twenty-feature ratio-model with 25 classes for hold-out method (Non-parametric estimation) for foreign material of CWRS wheat

Class (to) → (from) ↓	FM 3	FM 6	FM 9	FM 12	FM 15
FM 3	18	0	0	0	0
FM 6	0	16	2	0	0
FM 9	0	2	14	1	1
FM 12	1	1	2	13	1
FM 15	0	0	0	1	17

Table B4b. Confusion matrix of the twenty-feature ratio-model with 25 classes for the four layer BPNN for foreign material of CWRS wheat

Class (to) → (from) ↓	FM 3	FM 6	FM 9	FM 12	FM 15
FM 3	18	0	0	0	0
FM 6	1	16	1	0	0
FM 9	2	2	14	0	0
FM 12	1	1	0	14	2
FM 15	0	0	0	0	18

Table B4c. Confusion matrix of the twenty-feature slope-model with 25 classes for hold-out method (Non-parametric estimation) for foreign material of CWRS wheat

Class (to) → (from) ↓	FM 3	FM 6	FM 9	FM 12	FM 15
FM 3	16	2	0	0	0
FM 6	1	17	0	0	0
FM 9	0	1	16	1	0
FM 12	1	1	2	18	1
FM 15	0	0	0	1	17

Table B4d. Confusion matrix of the twenty-feature slope-model with 25 classes for the four layer BPNN for foreign material of CWRS wheat

Class (to) → (from) ↓	FM 3	FM 6	FM 9	FM 12	FM 15
FM 3	18	0	0	0	0
FM 6	1	16	1	0	0
FM 9	1	1	14	1	1
FM 12	1	0	0	15	2
FM 15	10	0	0	1	16

Table B4e. Confusion matrix of the twenty-feature slope ratio-model with 25 classes for hold-out method (Non-parametric estimation) for foreign material of CWRS wheat

Class (to) → (from) ↓	FM 3	FM 6	FM 9	FM 12	FM 15
FM 3	18	0	0	0	0
FM 6	0	16	2	0	0
FM 9	0	1	17	0	0
FM 12	0	0	0	17	1
FM 15	0	0	0	0	18

Table B4f. Confusion matrix of the twenty-feature slope ratio-model with 25 classes for the four layer BPNN for foreign material of CWRS wheat

Class (to) → (from) ↓	FM 3	FM 6	FM 9	FM 12	FM 15
FM 3	18	0	0	0	0
FM 6	0	18	0	0	0
FM 9	0	0	18	0	0
FM 12	1	0	0	17	0
FM 15	0	0	0	0	18