

**EVALUATION OF PHYSIOLOGICAL STATUS OF POTATO TUBERS USING
SPECTROSCOPIC AND HYPERSPECTRAL IMAGING SYSTEMS**

By

Ahmed Mustafa Rady

A DISSERTATION

Submitted to
Michigan State University
In partial fulfillment of the requirements
for the degree of

Biosystems Engineering - Doctor of Philosophy

2014

ABSTRACT

EVALUATION OF PHYSIOLOGICAL STATUS OF POTATO TUBERS USING SPECTROSCOPIC AND HYPSESPECTRAL IMAGING SYSTEMS

By

Ahmed Mustafa Rady

Potato is a major crop around the world with special importance given in developed countries to the French frying, and chipping industries. Quality attributes of potatoes dramatically influence final product conditions and consequently affect product marketability. Many research studies have been conducted to investigate the feasibility of measuring quality attributes and external and internal defects of potato tubers using rapid and/or noninvasive methods (spectroscopic, vision, and sonic). An extensive review has been conducted of nondestructive techniques that have been studied for assessing quality attributes of raw potatoes as well as chips and French fries. Such factors included specific gravity, dry matter, water content, carbohydrates, protein, defects, and diseases. In addition, systems for sorting tubers based on various quality characteristics have been discussed in detail. Also, commercial systems are available in the market for sorting and grading tubers based on different quality factors. However, more deep studies are needed to enhance rapid measurement performance and investigate more attributes that are important to growers and industry. The main objectives of this study were to investigate the potential of using spectroscopic as well as hyperspectral systems to evaluate processing-related constituents and parameters of two common potato cultivars, Frito Lay 1879 (FL) and Russet Norkotah (RN), using partial least squares regression (PLSR), and several types of artificial neural network (ANN) along with wavelengths selection techniques being interval partial least squares (IPLS), and genetic algorithm (GA). In addition, classification of tubers based on sugar levels has been conducted using linear discriminant analysis (LDA) functions, k-nearest neighbor (Knn), partial least squares discriminant analysis (PLSDA), feed forward artificial neural network (FFNN), and classifier fusion. The first study in the 2008 season was conducted to evaluate five constituents for both FL and RN using NIR transmittance, and VIS/NIR interactance modes as well as VIS/NIR hyperspectral systems for 0.5'' (12.7 mm) sliced samples and whole tubers. Results showed that the interactance mode yielded most of the best PLSR results. For primordium leaf counts, glucose, sucrose, specific gravity, and soluble solids, the optimum prediction models obtained from the interactance mode resulted in R (RPD) values of 0.95 (3.29), 0.90 (2.14), 0.81(1.63), 0.61(1.27), and 0.55(1.18) respectively for FL. For RN, the R(RPD) values were 0.90 (2.19), 0.95 (3.12), 0.63(1.30), 0.59(1.22), and 0.37(1.08) respectively. Slightly lower performance was achieved for whole tubers with optimal R(RPD) values FL in the case of primordium leaf counts, glucose, sucrose, and specific gravity of 0.89(2.22), 0.88(1.78), 0.81(1.64), and 0.37(1.06) respectively. The R(RPD) values for RN were 0.77(1.50), 0.79(1.60), 0.43(1.10), and 0.51(1.08) for primordium leaf counts, glucose, sucrose, and specific gravity. Soluble solids for whole tubers showed weaker correlation than above constituents.

Following preliminary results in the 2008 season, more concentration was given to glucose and sucrose as they significantly affect chip and French fry products quality. Also, based on preliminary results, transmittance mode was replaced by NIR reflectance mode. The second study was conducted in the 2009 and 2011 seasons using interactance, reflectance, and hyperspectral systems on the same cultivars and also using 0.5''(12.7 mm) sliced samples and whole tubers.

Results of prediction models using PLSR and ANN along with models using IPLS and GA as wavelength selection techniques demonstrated strong correlation for VIS/NIR hyperspectral systems in which only sliced samples were used. For glucose prediction models, R(RPD) values were as high as 0.81(1.70) and 0.97(3.66) for FL and RN and those values for the best sucrose prediction models were 0.58(1.23) and 0.38(1.0) for FL and RN. For VIS/NIR interactance mode, promising results for glucose prediction were shown for FL and RN. FL and RN yielded R(RPD) values of 0.92(2.35) and 0.95(3.02) respectively for sliced samples, and 0.85(1.92) for FL and 0.97(4.16) for RN in the case of whole tubers. Sucrose prediction models resulted in strong correlation with R(RPD) values as high as 0.95(3.29) and 0.78(1.57) for FL and RN for sliced samples, and 0.94(3.01) for FL and 0.94(2.82) for RN in the case of whole tubers. NIR reflectance showed auspicious performance for both cultivars. The best glucose prediction models yielded R(RPD) values for FL and RN as high as 0.96(3.47) and 0.97(4.21) in the case of sliced samples, and 0.82(1.78) and 0.98(4.84) for FL and RN in the case of whole tubers. Sucrose also showed high correlation for sliced samples with R(RPD) values of 0.96(3.89) and 0.97(3.92) for FL and RN, and those values for the whole tubers were 0.96(3.80) and 0.97(3.78) for FL and RN. In general, prediction models based on selected wavelengths showed similar or better performance compared to full wavelengths models, and it is worth stating that GA yielded higher numbers of selected variables (wavelengths) than IPLS; thus, the latter method was preferred as it often produced similar results compared to GA models.

Classification of potatoes based on sugar levels associated with the frying process showed encouraging results with the lowest classification error values of FL and RN obtained for glucose being 16% and 13%, for sliced samples, and 18% and 0% for whole tubers. In the case of sucrose models, error values in the case of sliced samples were 23% and 18%, and those values for whole tubers were 26% and 18% for FL and RN respectively. Such classification results for whole tubers demonstrated the feasibility of explaining more variation between samples when the data from interactance and reflectance modes was used, in the listed wavelengths ranges, and consequently applying both modes in an on-line system has the potential to enhance the sorting of tubers based on sugar levels.

This research demonstrated the feasibility of hyperspectral imaging systems as well as spectroscopic systems, in reflectance and interactance modes, in rapidly and accurately measuring some important constituents for potato growers and processing industries. Such results, especially for whole tubers, proved that there is a possibility for conducting an on-line sorting system based on sugar levels, or a hand-held device for rapid evaluation of quality either in field or during storage, to maintain potato tubers quality and accurately estimate the suitable time for harvesting or processing.

This work is greatly dedicated to my wife Mona Shaaban, and my kids Yusuf, Jana, and Omar

ACKNOWLEDGMENTS

Without the financial help from my country, Egypt, I wouldn't have been here to conduct my PhD. The governmental support from Egypt was the main reason to travel to USA, study at MSU, and become a Spartan.

During the course of my PhD program here in the Department of Biosystems and Agricultural Engineering, Michigan State University, I have been receiving advice and help from my major advisor Dr. Daniel E. Guyer. His help was effective in my study either inside or outside the academic field. I would also like to thank Dr. Renfu Lu for his guidance towards many technical contributions to my study. Also, I would like to express my deep gratitude to Dr. William Kirk for his support and advice during my program. My great appreciation also goes to Dr. Chai Lim for her technical support in data analysis.

I strongly express my thanks to my lab colleagues; Dr. Fernando Mendoza, Dr. Haiyan Cen, Dr. Irwin Donis Gonzalez, Dr. Akira Mizushima, and Mr. Benjamin Bailey for their big influence and help during my research. I learned many technical points in the field of nondestructive evaluation of food products from these friends and collaborators. Special thanks to Dr. Dennis Fulbright in the Department of Plant Pathology for his unlimited help in letting me use his lab. Also, I want to present my appreciation to Sara Stadt for her kind technical help during working in Dr. Fulbright's lab. Special appreciation to Rob Shafer; and Walther Farms, Three Rivers, MI, USA.

I greatly appreciate the scientific contribution in the BAE Department from Dr. Fred W-Bakker-Arkema, Dr. Ajit K. Srivastava, Dr. James Steffe, and Dr. Kirk Dolan. Also, I want to greatly thank Dr. Fathi Salem, and Dr. Lalita Upda in the ECE Department, and Dr. Anil K Jain

in the CSE Department for their influence towards building a strong background in the area of postharvest engineering of food and agricultural product..

I would like to present my great thanks to the Department of Biosystems and Agricultural Engineering, the Graduate School, and College of Agricultural and Natural Resources for providing financial help for travelling to a conference and support during the last year of my program.

The support before and during my PhD study from my wife, Mona Shaaban, and my mother-in-law was crucial in completing my PhD program. With the smiles on my kids, Yusuf, Jana, and Omar, I was able to continue and work hard on my research.

Last, and of big importance for me, I would like to give my unlimited thanks and appreciation to my mother and my sister for their great deal of patience and praying for me. And for my father's and brother's spirits that will be continuing to stay non-forgetful memories, I hope that I did what they always wish and I promise I will do my best to make them proud of me.

TABLE OF CONTENTS

LIST OF TABLES.....	xi
LIST OF FIGURES.....	xiv
KEY TO SYMBOLS OR ABBREVIATIONS.....	xvii
CHAPTER 1 INTRODUCTION.....	1
1.1 Economic Importance of Potatoes.....	1
1.2 Morphological Description of Potato Tuber.....	3
1.3 Overview.....	3
1.4 Objectives.....	4
CHAPTER 2 RAPID AND/OR NON-DESTRUCTIVE METHODS FOR QUALITY EVALUATION OF POTATOES: A REVIEW.....	6
2.1 Applications Related to Raw Tubers.....	7
2.1.1 Specific Gravity, Dry Matter, and Water Content.....	7
2.1.2 Carbohydrates and Protein Content.....	11
2.1.3 Defects and Diseases Detection of Potato Tubers.....	15
2.1.3.1 Spectroscopic-based methods.....	17
2.1.3.2 Imaging-based methods.....	18
2.1.3.3 Sonic-based methods.....	20
2.1.4 Systems for Non Destructive Sorting of Raw Potato Tubers.....	23
2.2 Applications of Non Destructive and/or Rapid Methods on Quality Evaluation for Potato Products.....	28
2.3 Commercial Sorting Systems for Potato Tubers, French Fries and Chips.....	33
2.4 Future Research.....	35
2.5 Summary.....	38
CHAPTER 3 THE POTENTIAL USE OF VISIBLE/NEAR INFRARED SPECTROSCOPY AND HYPERSPECTRAL IMAGING TO PREDICT PROCESSING-RELATED CONSTITUENTS OF POTATO TUBERS.....	40
3.1 Introduction.....	40
3.2 Materials and Methods.....	43
3.2.1 Sample Collection, Handling, and Treatments.....	43
3.2.2 Electronic Measurement.....	45
3.2.2.1 Sample preparation.....	45
3.2.2.2 VIS/NIR interactance mode.....	45
3.2.2.3 VIS/NIR hyperspectral mode.....	47
3.2.2.4 NIR transmittance mode.....	50
3.2.3 Constituent (Reference) Measurement.....	51
3.2.3.1 Measurement of glucose and sucrose.....	51
3.2.3.1.1 Extraction of juice.....	51
3.2.3.1.2 Chemical estimation of glucose and sucrose.....	52
3.2.3.2 Measurement of soluble solids.....	53
3.2.3.3 Measurement of specific gravity.....	53
3.2.3.4 Measurement of primordial leaf count.....	54
3.2.4 Partial Least Squares Regression (PLSR).....	54
3.2.4.1 Pretreatment of the spectra data.....	55
3.2.4.2 Pretreatment of the reference data.....	56
3.3 Results.....	59
3.3.1 Constituents' Distribution.....	59

3.3.2 Spectra for Different Modes.....	59
3.3.2.1 Interactance mode.....	59
3.3.2.2 Hyperspectral imaging mode.....	63
3.3.2.3 Transmittance mode.....	66
3.3.3 Partial Least Squares (PLSR) Results.....	68
3.3.3.1 Results for interactance mode.....	68
3.3.3.2. Results for hyperspectral reflectance mode.....	69
3.3.3.3 Results for transmittance mode.....	70
3.4 Discussion.....	71
3.5 Conclusions.....	72
CHAPTER 4 EVALUATION OF SUGAR CONTENT OF POTATOES USING HYPERSPSCTRAL IMAGING SYSTEMS.....	74
4.1 Introduction.....	74
4.2 Materials and Methods.....	76
4.2.1 Raw Material and Experimental Design.....	76
4.2.2. Constituent Measurement.....	80
4.2.2.1 Potato sample preparation.....	80
4.2.2.2 Wet chemistry basis measurements.....	80
4.2.2.2.1 Extraction of juice.....	80
4.2.2.2.2 Performing the chemical estimation of glucose and sucrose.....	81
4.2.3 VIS/NIR Hyperspectral Imaging Systems.....	81
4.2.4 Data Analysis Discussion and Approach.....	83
4.2.4.1 Definition and development of descriptive variables.....	83
4.2.4.1.1 Extracted mean spectra.....	83
4.2.4.1.2 Describing scattering profiles.....	85
4.2.4.2 Partial least squares regression (PLSR).....	86
4.2.4.3 Artificial neural network (ANN).....	86
4.2.4.4 Wavelength selection.....	92
4.2.4.5. Classification of potatoes based on sugar levels.....	93
4.3 Results and Discussions.....	94
4.3.1 Distribution of Glucose and Sucrose.....	94
4.3.2 Mean Reflectance Spectra (MRS).....	95
4.3.3 Curve Fitting Parameters.....	96
4.3.4 Note About Performance of the Hyperspectral System Used in the 2011 Season.....	99
4.3.5 Partial Least Squares Regression (PLSR) Results.....	99
4.3.6 Artificial Neural Network (ANN) Results.....	100
4.3.7. Variable Selection Results.....	102
4.3.8. Potato Classification Based on Sugar Levels.....	104
4.4 Conclusions.....	105
CHAPTER 5 UTILIZATION OF VISIBLE/NEAR-INFRARED SPECTROSCOPIC AND WAVELENGTH SELECTION METHODS IN SUGAR PREDICTION AND POTATOES CLASSIFICATION.....	107
5.1 Introduction.....	107
5.2 Materials and Methods.....	108
5.2.1 Raw Material and Experimental Design.....	108
5.2.2 Constituent Measurement.....	109
5.2.2.1 Potato sample preparation.....	109
5.2.2.2 Wet chemistry basis measurements.....	109
5.2.3 VIS/NIR Interactance System.....	109
5.2.4 Data Analysis Discussion and Approach.....	110
5.2.4.1 Data handling.....	110
5.2.4.2 Partial least squares regression (PLSR).....	110

5.2.4.3 Artificial neural network (ANN).....	110
5.2.4.4 Wavelength selection.....	110
5.2.4.5. Classification of potatoes based on sugar levels.....	112
5.3 Results and Discussions.....	113
5.3.1 Constituents Distribution.....	113
5.3.2 Spectra for Sliced Samples and Whole Tubers.....	113
5.3.3 Partial Least Squares Regression (PLSR) Results.....	118
5.3.3.1 Full and samples variables models.....	118
5.3.3.2 Selected variables-PLSR models.....	121
5.3.4 Artificial Neural Network (ANN) Results	126
5.3.4.1 Full and sampled variables models.....	126
5.3.4.2 Selected variables-ANN models.....	129
5.3.5 Results of Potatoes Classification Based on Sugar Levels and Selected Wavelengths.....	132
5.3.6 Results for 2009-2011 Combined Data.....	135
5.4 Conclusion.....	136
CHAPTER 6 RAPID EVALUATION OF PHYSIOLOGICAL STATUS OF POTATO TUBERS USING NEAR-INFRARED REFLECTANCE SPECTROSCOPIC METHODS.....	138
6.1 Introduction.....	138
6.2. Material and Methods.....	140
6.2.1. Constituent Measurement.....	140
6.2.1.1 Potato sample preparation.....	140
6.2.1.2 Wet chemistry basis measurements.....	140
6.2.2 NIR Reflectance System.....	140
6.2.3 Data Analysis Discussion and Approach.....	142
6.2.3.1 Data handling.....	142
6.2.3.2 Partial least squares regression (PLSR).....	143
6.2.3.3 Artificial neural network (ANN).....	143
6.2.3.4 Wavelength selection	143
6.2.3.5 Classification of potatoes based on sugar content.....	143
6.3 Results and Discussions.....	145
6.3.1 Constituents Distribution.....	145
6.3.2 Spectra for Sliced Samples and Whole Tubers.....	145
6.3.3 Partial Least Squares Regression (PLSR) Results.....	147
6.3.3.1 Full and sampled variables models.....	146
6.3.3.2 Selected variables-PLSR models.....	150
6.3.4 Artificial Neural Network (ANN) Results.....	155
6.3.4.1 Full and sampled variables models.....	155
6.3.4.2 Selected variables-ANN models.....	158
6.3.5 Results of Potatoes Classification Based on Sugar Levels and Selected Wavelengths.....	160
6.3.6 Results for 2009-2011 Combined Data.....	163
6.4 Conclusion.....	164
CHAPTER 7 INTEGRATING NIR REFLECTANCE, AND VIS/NIR INTERACTANCE SPECTROSCOPIC SYSTEMS DATA TO PREDICT PHYSIOLOGICAL STATUS OF POTATO TUBERS	165
7.1 Introduction.....	165
7.2 Materials and Methods.....	165
7.2.1 Raw Materials.....	165
7.2.2 Data Handling and Analysis.....	166
7.2.2.1 Data fusion.....	166
7.2.2.2 Data analysis.....	166
7.2.2.2.1 Partial least squares regression (PLSR).....	166
7.2.2.2.2 Artificial neural network (ANN).....	166
7.2.2.2.3 Classification of potatoes based on sugar levels.....	166
7.3 Results and Discussion.....	167

7.3.1 Partial Least Squares Regression (PLSR) Results	167
7.3.2 Artificial Neural Network (ANN) Results.....	169
7.3.3 Results for Classification of Potatoes Based On Sugar Levels.....	171
7.4 Conclusions.....	173
CHAPTER 8 OVERALL CONCLUSIONS AND FUTURE WORK.....	174
BIBLIOGRAPHY.....	182

LIST OF TABLES

Table 2.1.	Reported electronic techniques to estimate specific gravity, dry matter and water content for raw and non-processed potatoes.....	10
Table 2.2.	Reported electronic methods to estimate carbohydrates (starch, sugars), and protein content for raw and non-processed potatoes.....	15
Table 2.3.	Reported spectroscopic, imaging, and sonic methods to detect defects, and diseases for potato raw and non-processed potatoes.....	22
Table 2.4.	Reported electronic methods for sorting potato raw potato tubers using different quality attributes.....	25
Table 2.5.	Reported spectroscopic and imaging methods for assessing quality attributes for frozen French fries and potato chips.....	32
Table 2.6.	Commercial sorting and quality monitoring systems for raw potato tubers, French fries and chips.....	34
Table 3.1.	PLSR results for predicting some potato constituents using VIS/NIR interactance (sliced samples) for Frito Lay 1879 and Russet Norkotah cultivars.....	69
Table 3.2.	PLSR results for predicting some potato constituents using VIS/NIR interactance mode (whole tubers) for Frito Lay 1879 and Russet Norkotah cultivars.....	69
Table 3.3.	PLSR results for predicting some potato constituents using VIS/NIR hyperspectral imaging (sliced samples) for Frito Lay 1879 and Russet Norkotah cultivars.....	70
Table 3.4.	PLSR results for predicting some potato constituents using VIS/NIR hyperspectral imaging (whole tubers) for Frito Lay 1879 and Russet Norkotah cultivars	70
Table 3.5.	PLSR results for predicting some potato constituents using NIR transmittance (sliced samples) for Frito Lay 1879 and Russet Norkotah cultivars.....	71
Table 4.1.	Statistical summary of reference analysis resulted from wet chemistry for Frito Lay 1879 and Russet Norkotah cultivars.....	95
Table 4.2.	PLSR results of predicting glucose and sucrose using VIS/NIR hyperspectral imaging for sliced potato samples in the 2009 season using Frito Lay 1879 and Russet Norkotah cultivars.....	100
Table 4.3.	Results of prediction models to predict glucose and sucrose for sliced potato samples tested by VIS/NIR hyperspectral imaging and using RBFNN, RBFNNE, and FFNN in the 2009 season.....	101
Table 4.4.	Wavelength selection results using IPLS and GA in the case of glucose and sucrose for potato sliced samples tested VIS/NIR by hyperspectral imaging and in the 2009 season for Frito Lay 1879 and Russet Norkotah.....	102
Table 4.5.	PLSR results for predicting glucose and sucrose using VIS/NIR hyperspectral imaging and selected wavelengths obtained by IPLS and GA for sliced samples in the 2009 season for Frito Lay 1879 and Russet Norkotah cultivars.....	103

Table 4.6.	Artificial neural network results for predicting glucose and sucrose using VIS/NIR hyperspectral imaging and selected wavelengths obtained by IPLS and GA for sliced samples in the 2009 season for Frito Lay 1879 and Russet Norkotah cultivars.....	103
Table 4.7.	Numbers of samples in each class based on glucose and sucrose levels for the 2009 season in the case of Frito Lay1879 and Russet Norkotah cultivars.....	104
Table 4.8.	Classification results of sliced samples based on glucose and sucrose levels for the 2009 season using VIS/NIR hyperspectral imaging for Frito Lay1879 and Russet Norkotah cultivars.....	104
Table 5.1.	PLSR results for predicting glucose and sucrose for sliced samples and whole tubers using VIS/NIR interactance and using full (2701) and sampled wavelengths in the 2009 and 2011 seasons for Frito Lay 1879 and Russet Norkotah cultivars.....	120
Table 5.2.	PLSR results for predicting glucose and sucrose for sliced samples and whole tubers using selected wavelengths obtained by IPLS and GA (from sampled wavelengths) and VIS/NIR interactance in the 2009 and 2011seasons for Frito Lay 1879 and Russet Norkotah cultivars.....	123
Table 5.3.	Selected wavelengths for predicting glucose and sucrose for sliced samples and whole tubers using IPLS and GA methods (from sampled wavelengths) and VIS/NIR interactance in the 2009 and 2011 seasons for Frito Lay 1879 and Russet Norkotah cultivars.....	125
Table 5.4.	ANN results for predicting glucose and sucrose for sliced samples and whole tubers using VIS/NIR interactance and using full (2701) and sampled wavelengths in the 2009 and 2011 seasons for Frito Lay 1879 and Russet Norkotah cultivars.....	128
Table 5.5.	ANN results for predicting glucose and sucrose for sliced samples and whole tubers using selected wavelengths obtained by IPLS and GA (from sampled wavelengths) and VIS/NIR interactance in the 2009 and 2011seasons for Frito Lay 1879 and Russet Norkotah cultivars.....	131
Table 5.6.	Number of samples in each class based on glucose and sucrose levels, obtained from wet chemistry, for sliced samples and whole tubers in the 2009 and 2011 seasons for Frito Lay 1879 and Russet Norkotah cultivars.....	132
Table 5.7.	Classification results of sliced samples and whole tubers of Frito Lay 1879 and Russet Norkotah cultivars based on glucose and sucrose levels and using multiple classification techniques and VIS/NIR interactance in the 2009 and 2011 seasons.....	134
Table 5.8.	PLSR results for predicting glucose and sucrose for sliced samples and whole tubers using VIS/NIR interactance for Frito Lay 1879 and Russet Norkotah cultivars using 2009 and 2011 combined data.....	135
Table 5.9.	ANN results for predicting glucose and sucrose for sliced samples and whole tubers using VIS/NIR interactance for Frito Lay 1879 and Russet Norkotah cultivars using 2009 and 2011 combined data.....	136
Table 6.1.	PLSR results for predicting glucose and sucrose for sliced samples and whole tubers using NIR reflectance and using full (784) and sampled wavelengths in the 2009 and 2011 seasons for Frito Lay 1879 and Russet Norkotah cultivars.....	149

Table 6.2.	PLSR results for predicting glucose and sucrose for sliced samples and whole tubers using selected wavelengths obtained by IPLS and GA (from sampled wavelengths) and NIR reflectance in the 2009 and 2011 seasons for Frito Lay 1879 and Russet Norkotah cultivars...	153
Table 6.3.	Selected wavelengths for predicting glucose and sucrose for sliced samples and whole tubers using IPLS and GA methods (from sampled wavelengths) and NIR reflectance in the 2009 and 2011 seasons for Frito Lay 1879 and Russet Norkotah cultivars.....	154
Table 6.4.	ANN results for predicting glucose and sucrose for sliced samples and whole tubers using NIR reflectance and using full (784) and sampled wavelengths in the 2009 and 2011 seasons for Frito Lay 1879 and Russet Norkotah cultivars.....	157
Table 6.5.	ANN results for predicting glucose and sucrose for sliced samples and whole tubers using selected wavelengths obtained by IPLS and GA (from sampled wavelengths) and NIR reflectance in the 2009 and 2011 seasons for Frito Lay 1879 and Russet Norkotah cultivars...	159
Table 6.6.	Classification results of sliced samples and whole tubers of Frito Lay 1879 and Russet Norkotah cultivars based on glucose and sucrose levels and using multiple classification techniques and NIR reflectance in the 2009 and 2011 seasons.....	162
Table 6.7.	PLSR results for predicting glucose and sucrose for sliced samples and whole tubers using NIR reflectance for Frito Lay 1879 and Russet Norkotah cultivars using 2009 and 2011 combined data.....	163
Table 6.8.	ANN results for predicting glucose and sucrose for sliced samples and whole tubers using NIR reflectance for Frito Lay 1879 and Russet Norkotah cultivars using 2009 and 2011 combined data.....	163
Table 7.1.	Summary of the best prediction models using PLSR for glucose and sucrose using VIS/NIR interactance and NIR reflectance individual modes for sliced samples and whole tubers for Frito Lay 1879 and Russet Norkotah cultivars.....	168
Table 7.2.	PLSR results for predicting glucose and sucrose using fused data from VIS/NIR interactance and NIR reflectance systems for sliced samples and whole tubers for Frito Lay 1879 and Russet Norkotah cultivars in the 2009 and 2011 seasons.....	168
Table 7.3.	Summary of the best prediction models using ANN for glucose and sucrose using VIS/NIR interactance and NIR reflectance individual modes for sliced samples and whole tubers for Frito Lay 1879 and Russet Norkotah cultivars.....	169
Table 7.4.	ANN results for predicting glucose and sucrose using fused data from VIS/NIR interactance and NIR reflectance systems for sliced samples and whole tubers for Frito Lay 1879 and Russet Norkotah cultivars in the 2009 and 2011 seasons.....	170
Table 7.5.	Summary of the best classification results based on glucose and sucrose levels using VIS/NIR interactance and NIR reflectance individual modes for sliced samples and whole tubers for Frito Lay 1879 and Russet Norkotah cultivars.....	171
Table 7.6.	Classification results of sliced samples and whole tubers based on glucose and sucrose levels for Frito Lay 1879 and Russet Norkotah cultivars using multiple classification techniques and VIS/NIR interactance and NIR reflectance combined data sets in the 2009 and 2011 seasons.....	172

LIST OF FIGURES

Figure 2.1.	Breakdown of reviewed technology used in potato postharvest and quality assurance of some potato products.....	7
Figure 3.1.	Flow chart of the experimental design to assess physiological status of potato tubers using visible/near infrared spectroscopy and hyperspectral imaging for Frito Lay 1879 and Russet Norkotah cultivars.....	44
Figure 3.2.	a. Schematic representation of VIS/NIR interactance mode used to predict constituents for two potato cultivars, b. Light path representation, c. End view of probe.....	46
Figure 3.3.	a. Schematic representation of VIS/NIR hyperspectral reflectance mode used to predict constituents for two potato cultivars, b. Light scattering in sample and scanning configuration.....	48
Figure 3.4.	a. Hyperspectral scattering image, with different colors representing light intensity of a potato slice, b. Spectral profiles from different spatial locations represented by different colors, c. Spectral profiles from different wavelengths represented by different colors.....	49
Figure 3.5.	a. Schematic representation of NIR transmittance mode configuration and system components, b. Light path representation with scattering in the sample and the detected transmitting light.....	51
Figure 3.6.	Flow chart of preprocessing methods used to pretreat spectra, a, and reference, b, data before building calibration and then prediction models using PLSR with cross validation to predict constituents for two potato cultivars.....	58
Figure 3.7.	Data distributions of the physiological variables measured; a. Glucose concentration % (note change in range of the values for the two cultivar types), b. Sucrose concentration (%), Primordial leaf count (number of leaves per sprout), d. Specific gravity (g/cm ³), e. Soluble solids (Brix scale).....	60
Figure 3.8.	Mean relative interactance for two sugar groupings for sliced samples, a. Frito Lay 1879: glucose, b. Russet Norkotah: glucose, c. Frito Lay 1879: sucrose, and d. Russet Norkotah: sucrose.....	61
Figure 3.9.	Mean relative interactance for two sugar groupings for whole tubers, a. Frito Lay 1879: glucose, b. Russet Norkotah: glucose, c. Frito Lay 1879: sucrose, and d. Russet Norkotah: sucrose.....	62
Figure 3.10.	Mean relative reflectance for two sugar groupings for sliced samples, a. Frito Lay 1879: glucose, b. Russet Norkotah: glucose, c. Frito Lay 1879: sucrose, and d. Russet Norkotah: sucrose.....	64
Figure 3.11.	Mean relative reflectance for two groupings for whole tubers, a. Frito Lay 1879: glucose, b. Russet Norkotah: glucose, c. Frito Lay 1879: sucrose, and d. Russet Norkotah: sucrose.....	65
Figure 3.12.	Relative transmittance for two sugar groupings for sliced samples, a. Frito Lay 1879: glucose, b. Russet Norkotah: glucose, c. Frito Lay 1879: sucrose, and d. Russet Norkotah: sucrose.....	67
Figure 4.1.	Flow chart of the experimental design to assess physiological status of potato tubers using VIS/NIR hyperspectral imaging for Frito Lay 1879 and Russet Norkotah cultivars in the 2009 season.....	77

Figure 4.2.	Flow chart of the experimental design to assess physiological status of potato tubers using VIS/NIR hyperspectral imaging for Frito Lay 1879 and Russet Norkotah cultivars in the 2011 season.....	79
Figure 4.3.	a. Hyperspectral imaging Optical Properties Analyzer (OPA) used in the 2011 season. b, Schematic of OPA.....	82
Figure 4.4.	a. An example of an image obtained for each slice sample, b. Sample of spectra at different wavelengths, c. Sample of average spectrum for one image.....	84
Figure 4.5.	Figure 4.5. Decaying portion of original spatial scattering profiles for selected sliced samples of Frito Lay 1879 cultivar at 698.7 nm in the 2009 season.....	86
Figure 4.6.	Schematic representation of RBFNN (after Haykin, 2009).....	89
Figure 4.7.	Schematic representation of FFNN (after Varmuza and Flizmoser, 2007).....	91
Figure 4.8.	Relative mean reflectance for a. Frito Lay 1879, b. Russet Norkotah, for the 2009 season, and relative mean reflectance for c. Frito Lay 1879, and d. Russet Norkotah, for the 2011 season.....	96
Figure 4.9.	Relative parameter a_{wi} for a. Frito Lay 1879, b. Russet Norkotah, for the 2009 season, and relative parameter a_{wi} for c. Frito Lay 1879, and d. Russet Norkotah, for the 2011 season.....	97
Figure 4.10:	Relative parameter b_{wi} for a. Frito Lay 1879, b. Russet Norkotah, for the 2009 season, and relative parameter b_{wi} for c. Frito Lay 1879, and d. Russet Norkotah, for the 2011 season.....	98
Figure 4.11.	Relationship between measured and predicted glucose values for sliced samples using full wavelengths for a) Frito Lay1879 and b) Russet Norkotah cultivars in the 2009 season using PLSR as indicated in table 4.2.....	100
Figure 4.12:	Relationship between measured and predicted glucose values for sliced samples using full wavelengths for a) Frito Lay1879 and b) Russet Norkotah in the 2009 season using ANN as indicated in table 4.3.....	101
Figure 5.1.	A schematic diagram of data handling and analysis for data obtained using VIS/NIR interactance spectroscopy to obtain prediction models of glucose and sucrose and for classification of Frito Lay1879 and Russet Norkotah based on sugar levels.....	111
Figure 5.2.	Distribution of glucose and sucrose (%FW) for Frito Lay 1879 and Russet Norkotah from wet chemistry in a) 2009, and b) 2011 seasons. Note: scale change on RN glucose for display purpose.....	114
Figure 5.3.	Relative interactance of the 2009 season data for sliced samples a. Frito Lay 1879, b. Russet Norkotah, and relative interactance for whole tubers for c. Frito Lay 1879, and d. Russet Norkotah.....	115
Figure 5.4.	Relative interactance of the 2011 season data for sliced samples a. Frito Lay 1879, b. Russet Norkotah, and relative interactance for whole tubers for c. Frito Lay 1879, and d. Russet Norkotah.....	116
Figure 5.5.	Mean of log (1/relative interactance) of the 2009 season data for Frito Lay 1879 and Russet Norkotah for: a. Sliced samples, b. Whole tubers	118

Figure 5.6.	Best relationships between wet chemistry based and PLSR predicted constituents for Frito Lay 1879 and Russet Norkotah in the 2009 season for a) Glucose for sliced samples, b) Glucose for whole tubers, c) Sucrose for sliced samples, and d) Sucrose for whole tubers.....	124
Figure 5.7.	Schematic representation of the selected wavelengths, using VIS/NIR interactance mode and IPLS, associated with the best PLSR models of glucose and sucrose in the 2009 and 2011 seasons for sliced samples and whole tubers for a) Frito Lay 1879, b) Russet Norkotah.....	126
Figure.6.1.	Schematic representation of NIR diffuse reflectance mode and a clearer view of sample setting	142
Figure 6.2.	A schematic diagram of data handling and analysis for NIR reflectance experiments to obtain prediction models of glucose and sucrose for two potato cultivars.....	144
Figure 6.3.	Signals of absorbance ($\log(1/\text{relative reflectance})$) for the 2009 season data for sliced samples a. Frito Lay 1879, b. Russet Norkotah, and for whole tubers for c. Frito Lay 1879, and d. Russet Norkotah.....	146
Figure 6.4.	Signals of absorbance ($\log(1/\text{relative reflectance})$) for the 2011 season data for sliced samples a. Frito Lay 1879, b. Russet Norkotah, and for whole tubers for c. Frito Lay 1879, and d. Russet Norkotah.....	147
Figure 6.5.	Best prediction models based on selected wavelengths and PLSR predicted constituents in the 2011 season for Frito Lay 1879 and Russet Norkotah cultivars for a) Glucose for sliced samples, b) Glucose for whole tubers, c) Sucrose for sliced samples, and d) Sucrose for whole tubers.....	155

KEY TO SYMBOLS FOR ABBREVIATIONS

ANN	Artificial neural network
CCD	Charge-coupled device
FFNN	Feed forward neural networks with back propagation
FL	Frito Lay1879
FLGL	Frito Lay1879 glucose
FLSU	Frito Lay1879 sucrose
FWHM	Full width half maximum
GA	Genetic algorithm
HIS	Hyperspectral imaging system
IPLS	Interval partial least squares
LV	Latent variables
MRS	Mean reflectance spectra
NEWGRNN	Generalized radial basis function neural networks
PLSR	Partial least squares regression
R or r	Correlation coefficient
RBFINN	Radial basis function neural networks
RBFINNE	Exact design radial basis function neural networks
RMSE	Root mean square of error
RMSEC _{cv}	Root mean square error of calibration using cross validation
RMSEP	Root mean square error of prediction
RN	Russet Norkotah
RNGL	Russet Norkotah glucose

RNSU	Russet Norkotah sucrose
RPD	Ratio of performance to deviation
R_{cal}	Correlation coefficient of calibration model
R_{pred}	Coefficient of correlation of prediction model
R_{test}	Correlation coefficient of validation model
R_{train}	Correlation coefficient of training model
$SeCV_{train}$	Standard error of training model using cross validation
Se_{test}	Standard error of test set

CHAPTER 1 INTRODUCTION

1.1 Economic Importance of Potatoes

Potatoes (*Solanum tuberosum*) rank as the fifth highest produced commodity used for human consumption (after: sugar cane, maize, wheat, rice), with the world production of 324.2 million metric ton (FAOSTAT, 2012). Potato is a common source of carbohydrate with a diverse set of uses. In North America, Europe, and Australia, the majority of potatoes are processed as chips (crisps), French fries, dehydrated, canned, mashed, diced, etc. For example, in U.S.A, which produces 18.3 million metric ton, only one-third is consumed as fresh product; approximately, 60% is consumed as processed products, and 6% is used as seeds (FAOSTAT, 2012). However, other countries, especially the developing ones, consume the majority of potatoes as fresh due to the living and income style. The modern life style promotes high quality foodstuff products in either home-prepared or fast food which increases the need for efficient, rapid, cost effective and easy to use devices and systems to assure that final product, processed or fresh, meets the required quality. There has been a considerable change in the trend of potato product consumption in the US since 1960. Almost 38.10 Kg per capita fresh tuber was consumed in 1960, and that number decreased to 19.05 Kg per capita by 2005. However, frozen potato fries and chips consumption increased from 8.16 to 31.75 Kg per capita from 1960 to 2005. Moreover, the value of US exports of chips and frozen French fries increased from \$610 million in 2006 to more \$810 million in 2010 (Bohl and Johnson, 2010). On the other hand, the USA was the highest country outside the European Union in importing French fries with a quantitative increase of 28% from 2000 to 2010. Other developed countries such as: Japan, Canada, and Australia also showed a considerable increase in imports of French fries in the same

time span of 22%, 435%, and 558% respectively. A similar trend was noticed for some of the developing countries such as Brazil (229 %), Mexico (177%), and China (241%) (Faulkner, 2012). Given the previous statistics, one can conclude that there is a tremendous growth of processed potato products in many countries (developed and developing) that requires reliable, accurate, rapid, and reproducible systems to maintain quality aspects of tubers and final products. The more importance given to preserving the high quality of potatoes before processing, the higher marketability of products, and consequently more benefit for growers, processors, and consumers.

Potato global trade has shown increase through the last three decades. International potato trade has doubled from 1986 (<10 MT) to 2006 with a quadrupled value in the same time range reaching \$9.6 billion in 2007 with majority of which being processed (Chrome et al, 2010). Moreover, In USA, \$9 billion was spent on potato chips in 2011. Such statistics show the economic value of potato products in US as well as in the world.

Some chemical constituents and physical properties in potato tubers determine their end use for either the processed industry or as fresh, or prevent the use of tubers if the levels of these parameters are beyond the suggested thresholds. These constituents are, but not limited to, specific gravity, carbohydrate, protein, vitamins, glycoalkaloids, minerals, flesh and skin color, carotenoids, and anthocyanins. Other aspects which determine quality and potential use are the external or internal defects such as greening, bruises, enzymatic browning, non-enzymatic browning, and other physiological disorders, (Storey, 2007). Additionally, the accurate estimation of optimal harvest time is critical for potato tubers as it strongly affects quality of the harvested tubers.

1.2 Morphological Description of Potato Tuber

A potato tuber is a modified stem with leaves and axillary buds that are reduced and poorly developed. In addition, a potato tuber has shortened internodes and a stem axis that expands radially. The end of the tuber attached to the stolon is called the heel, or stem end, while the other end is called the rose end or stolon apex. A potato tuber is considered a third type of stem in a potato plant as there are the regular above-ground stems, and the stolon which is the under-ground stem. Moreover, potato tubers are considered as the swollen parts of the stolon which is the rhizome of the potato plant. Stolons are diageotropic shoots or stems with elongated internodes and rudimentary leaves. Stolons are grown from the basal stem nodes below the soil surface. Stolonization mostly starts after 15 days from planting and at the nodes closer to the seed tuber and then progress acropetally. Earlier stolons grow faster and become longer than later ones and the number of stolons increases with time. Tuber formation can be thought as of the result of two operations: stolon formation, or stolonization, and tuberization of the stolon tip. Tubers are formed after 25 days from planting in most potato cultivars. Tuberization starts before all stolons are formed, and it occurs first in the lower stolons and results in dominant tubers in terms of weight over those formed later. Sugar in potato tubers are either monosaccharaides (glucose and fructose) with portions in the tuber of 0.15-1.5% of FWT for either sugar, or disaccharides that is the sucrose sugar and its levels are 0.4-6.6%. Sugar levels depend on cultivar, preharvest treatments, storage temperature and period.

1.3 Overview

In this study, the experiments were conducted on three seasons. In the first season, 2008, VIS/NIR spectroscopic system in the interactance mode, NIR transmittance, and VIS/NIR hyperspectral imaging were used to study the prediction of glucose, sucrose, specific gravity,

soluble solids, and primordial leaf count for Frito Lay 1879, and Russet Norkotah potato cultivars. Whole tubers and 0.5'' (12.7 mm) sliced samples were tested to build calibration and prediction models using PLSR. Based on the results obtained in the 2008 season, electronic measurements were narrowed to glucose and sucrose in the 2009 and 2011 seasons. Moreover, the transmittance mode has been replaced by the reflectance mode in the same wavelength range. Additionally, in the 2009 and 2011 seasons, an artificial neural network (ANN) technique was also used for building prediction models, and the most influential wavelengths were identified using IPLS and GA. Finally, the classification of potato tubers and sliced samples has been studied based on glucose and sucrose.

1.4 Objectives

This research was conducted to study the feasibility of using spectroscopic and hyperspectral imaging systems to evaluate some constituents of potato tubers for some cultivars used in processing, table, or seed industries. Consequently, the particular objectives of work were:

- 1) Determine calibration and prediction models for glucose and sucrose for potato tubers using different regression methods on spectroscopic and hyperspectral imaging data against traditional chemistry-based measurements.
- 2) Identify the most effective wavelengths related to glucose and sucrose prediction in potato tubers.
- 3) Determine whether combined data from different spectroscopic and hyperspectral systems (sensor fusion) can improve prediction models of glucose and sucrose for potato tubers.

- 4) Study the potential of using spectroscopic and hyperspectral imaging systems in potato classification based on sugar levels associated with frying.

CHAPTER 2 RAPID AND/OR NON-DESTRUCTIVE METHODS FOR QUALITY EVALUATION OF POTATOES: A REVIEW

(Rady, A.M., Guyer, D.E. 2014. Rapid and/or Non-Destructive Quality Evaluation Methods for Potatoes: A Review. Computers and Electronics in Agriculture (in review))

There are many rapid techniques which have been used in attempts to evaluate the physiological status of potato tubers as well as to test the quality attributes of finished potato products. These systems range in basic operation theory and they include: traditional imaging systems (CCD cameras, multispectral imaging, X-ray, magnetic resonance imaging (MRI)), spectroscopic systems (UV, visual, near, and mid infra-red systems), hyperspectral imaging systems, and ultrasonic systems. The applications of such systems for raw tubers include predicting of chemical constituents and physical characteristics (dry matter, specific gravity, carbohydrate, and water content), detecting of defects and diseases, and electronic-based sorting. Other applications address automated quality evaluation of potato products (chips, French fries). While there have been a significant number of studies regarding the application of rapid estimation of quality attributes of raw potato tubers and processed products, no study was conducted to summarize such different approaches. In addition to discussing the above systems overall, this review aims to present some of the commercial systems that exist for the potato industry. Fig.2.1 depicts the applications of non-destructive methods for postharvest potatoes and potato products reviewed in this study.

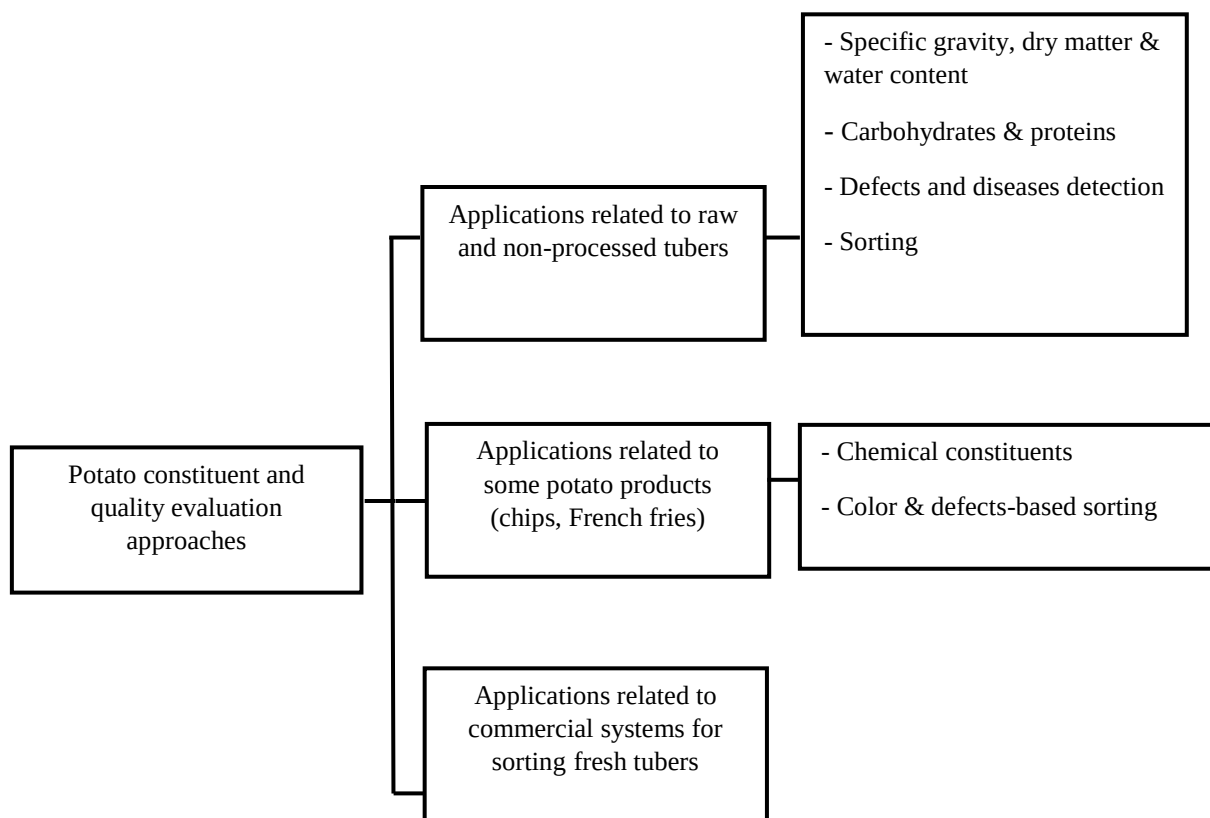


Figure 2.1. Breakdown of reviewed technology used in potato postharvest and quality assurance of some potato products.

2.1 Applications Related to Raw Tubers

These applications vary based on the material being tested: whole tubers, sliced samples, or any other non-cooked forms. The applications are mostly related to measuring quality assurance attributes right before harvesting, after harvest, or in the storage facilities, and the use of such tubers might be fresh, or preparing for processing.

2.1.1 Specific Gravity, Dry Matter, and Water Content

Being one of the most important factors for assessing potato tubers for processing, specific gravity (SG) is strongly affected by: environmental factors (weather, soil type), variety, and production operations: seed management, plant density, nutrient management, irrigation, tuber growth period, disease management, vine killing, and harvest management (Stark and

Love, 2003). Higher SG results in more output of chip, French fry, and dehydrated products. Levels for SG in potato tubers are crucial for processing. A SG level of 1.08 or higher is preferred for chipping (Gould, 1995a; Stark et al., 2003). Literature confirmed the positive correlation, with linear relationship, between SG and dry matter (DM) with many equations found to obtain the SG from dry matter (Woodbury and Weinheimer, 1964; Houghland, 1966; Agle and Woobery, 1968; Willson and Lindsay, 1969; Schippers, 1976; Simmonds, 1977). Therefore, the SG is generally and extensively used as a stick measure of dry matter and to estimate the suitability of tubers destined to processing. Moreover, DM is commonly correlated to the texture quality of raw and cooked tubers which is evaluated by sensory-related tests (Tarn et al., 1992). DM content is about 18 to 26% for most potato cultivars dedicated for commercial use (Burton, 1989). Desirable levels of DM for processing depend on the use of potatoes. DM ratios of 20-24%, 22-24, and >21% are preferred for French fries, chipping and dehydrated industries respectively.

DM distribution inside tubers was studied by many researchers (Glynne and Jackson, 1919; Johnston et al., 1968; Pritchard and Scanlon, 1997; Gaze et al., 1998). It was shown that DM is more concentrated in the storage parenchyma between the cortex and the vascular ring, and longitudinally decreases in towards the pith. There have been two common methods to estimate specific gravity, the first one based on the weight in air vs. weight in water relationship, and the other is using a hydrometer. Both methods, however, are time consuming, depend on human proficiency level, and do not cope with on-line sorting applications based on SG.

Therefore, several rapid techniques, most of which are spectroscopic-based systems, have been tested to estimate either the specific gravity or dry matter as illustrated in table 2.1. Model accuracy for spectroscopic systems is usually judged using root mean error of calibration

(RMSEC), root mean error of calibration using cross validation (RMSEC_{cv}), or prediction (RMSEP), coefficient of correlation (R), and/or the ratio of the standard deviation of reference variable to RMSEP or RMSEC which is abbreviated as RPD. Values of R for prediction or validation models, and RMSEP or RMSEC_{cv} are listed in this review study; otherwise R values for calibration models are listed. The sign (?) was used in table 2.1 and subsequent tables in the case of the non-availability of model strength descriptive values. Among the varying types of electronic systems used for the evaluation of DM in potatoes and applied on various sampling techniques, studies conducted by Hartmann and Büning-Bfaue, 1998; Haase (2004 and 2011), on homogenized, mashed, and ground samples, respectively, NIR or VIS/NIR reflectance (1100-2500 nm, 300-2500 nm, and 850-2500 nm) yielded the best prediction performance (RMSEP= 0.19%, 0.568%, and 0.42%). Generally, it was shown that NIR radiation intensity inside fruit tissue decreases in an exponential trend with depth (Lammertyn et al., 2000; Fraser et al., 2000). Consequently, having relatively lower performance for whole tubers, and slices, for estimating chemical constituents can be understandable. Moreover, skin is a factor resulting in dispersing, interfering, and weakening of detected signals, and mostly yields lower correlation between spectra and chemical compounds inside the tissue (Fraser et al., 2003). Therefore, and based on the DM distribution inside potato tubers, reflectance and interactance modes generally yielded better correlation than transmittance mode. However, sampling methods applied on such studies are not suitable for on-line sorting. SG prediction models showed the same performance as DM, between the three spectroscopic modes, which probably is a result of the fact that SG is a direct indication of DM or the solids inside the tuber.

Water content (WC) is also an important factor of potato tubers as it is inversely proportional with DM, SG, and starch content. Thus, it's desirable to keep WC in potato tubers at

levels that protect tubers from water loss and shrinkage without any excess that reduces tubers' suitability for processing. Water absorption peaks in the NIR range are located at 970, 1200, and 1450 nm (Workman, and Weyer, 2008). In some cases, however, some interfering might occur between water and other constituents' absorption when using a broad wavelength range. Consequently, table 2.1 shows that relatively low RMSEP values were obtained for WC estimation when using narrower wavelength range as conducted by Qiao et al., 2005 (RMSEP= 0.14%) compared with RMSEP values of 6.414%, 4.791%, 1.761%, and 0.387% obtained by Singh et al., 2004.

Table 2.1. Reported electronic techniques to estimate specific gravity, dry matter and water content for raw and non-processed potatoes.

Mode(spectral range)	Parameter	Tested material	R%(RMSEP)	Reference
NIR transmittance (800-1000 nm)	DM	Whole tubers Thin slices(2.54 cm) Thick slices(4-6 cm)	92(1.52%) (no test set) 97(?) 95(1.69%)	(Dull et al., 1989)
NIR reflectance (1100-2500 nm)	DM	Homogenized	97(0.19%) (no test set)	(Hartmann and Büning-Pfaue, 1998)
NIR reflectance (770-2498 nm)	SG DM	Cylindrical	87(0.007) 88(1.3%)	Scanlon et al., 1999)
NMR (low field)	DM	Slices Raw Boiled	?(?) ?(?) ?(?)	(Thybo et al., 2000)
NMR (low field)	DM	Slices	?(?)	(Thygesen et al., 2001)
VIS/NIR transmittance (530-1100)	SG	Whole tubers Punctured tubers	85(0.002) 82(0.002)	(Kang et al., 2003)
NMR (low field)/ MRI	DM	Slice Cylindrical	?(?) ?(?)	(Thybo et al., 2003)
VIS/NIR reflectance (400-2500 nm)	DM	Mashed	98(0.533%)	(Haase, 2004)
VIS/NIR reflectance (300-2500 nm)	WC	slab samples (6x4x0.3 cm) Without skin(400-1750) Without skin(700-900,1000-1100,1250-1600) With skin(400-1750) With skin(850-900,1100-1200,1400-1500)	99(6.414%) 99(4.791%) 99(1.761%) 99(0.387%)	(Singh et al., 2004)
NIR transmittance	DM	Whole tuber	79(1.04%) (no test set)	(Walsh et al., 2004)
VIS/NIR interactance (400-1100 nm)	SG	Whole tubers	90(0.004) (no test set)	(Chen et al., 2005)
Hyperspectral imaging (934-997 nm)	WC	Whole tuber	88(0.014%)	(Qiao et al., 2005)
NIR interactance (750-950 nm)	DM	Peeled Slices Slices(moving)	95(1.13%) 93(1.08%) 90(1.08%)	(Subedi and Walsh, 2009)
NIR reflectance (850-2500 nm)	DM	Ground samples	99(0.42%)	(Haase, 2011)
1 D VIS/NIR interactance (449-1040 nm) 2 D NIR interactance (760-1040 nm)	DM	Whole tubers (unpeeled)	97(0.91%) (no test set) 91(1.68%)	(Helgerud et al., 2012)

2.1.2 Carbohydrates and Protein Content

Potato is known as a good source of carbohydrates in comparison with grains. Carbohydrate concentrations, as well as chemical constituents, depend on variety, soil type, cultural practice, maturity stage, diseases, and storage conditions (Rama, and Narasimham, 2003). Total carbohydrates significantly differs between the raw potato (18.5% FW), and dried potato (74.3% weight) which exceeds or is close to the same value for other carbohydrate sources: rice (80.2% FW), wheat (70.9% FW), sweet potato (27.4% FW), yam (24.2% FW), and cassava (35.2% FW) (Woolfe, 1987). McCay et al. (1975), stated that the number of calories obtained from one medium size potato tuber is the same as that obtained from an apple or a banana. Carbohydrates in potato tubers include: starch, sugars, cellulose, hemicellulose, and other polysaccharides. Starch is the major component in potato carbohydrates accounting for 60-80% of the dry matter (Kadam, et al.,1991). There are two main types of starch in potato tubers: amylose (linear chain of glucose molecules linked by 1,4-glycosidic bonds) that account for 20% of the tuber starch and the rest is amylopectin in which the glucose chains are also branched by 1,6 glycosidic bonds (Storey, 2007).

Starch concentration in potato tubers starts with low levels after tuber initiation with an increase during buckling, and reaches its maximum value at the start of the senescence process. Starch then decreases with the time of vine killing in a similar trend to the specific gravity accumulation process (Stark and Love, 2003). Starch is shown to positively correlate with specific gravity and dry matter. Tubers with starch content of 13% or higher are acceptable for processing (Stark et al., 2003). The common method to determine total starch in potato tubers is the enzymatic hydrolysis in which the starch is completely converted into D-glucose using specific enzymes. In addition of being a destructive method requiring preparation time, the

enzymatic method has some possible drawbacks including the interfering of other enzymes that leads to higher or lower total starch levels than the actual content (BeMiller, 2003).

Monitoring starch content in potatoes, as shown in table 2.2, using spectroscopic systems was feasible with most of tests conducted on mashed or ground tubers, and resulting in relatively low RMSEP values (0.651%, and 0.740% by Haase 2004; and Haase, 2011) compared to limits recommended for tuber processing. The relatively high content of starch in potatoes, compared to other constituents, and the broad distribution inside the tuber (in cortex, vascular ring, and parenchyma) resulted in strong correlation with NIR, or VIS/NIR spectroscopic systems. Some studies showed standard error of prediction (SEP) rather than RMSEP. Studies with a separate test set (Haase, 2004; Haase, 2006; and Haase, 2011) resulted in higher error values than that with only validation set which is statistically expectable (Hartmann and Büning-Pfaue, 1998).

The main reducing sugars in potato tubers are: glucose (0.15-1.5%FW), and fructose (0.15-1.5%FW) which are reducing sugars, and sucrose (0.4-6.6%FW), (Storey, 2007). Sugar level varies with variety, and low sugar varieties are usually dedicated for processing (Liu et al., 2009). There are different scenarios for sugar formation in potato tubers; sucrose is usually formed during the photosynthesis process, it then is enzymatically divided into glucose and fructose. Fructose is converted into glucose and the glucose forms the starch molecules (Stark and Love, 2003). Traditional methods of sugar measurement include: HPLC (high performance liquid chromatography), HPAEC (high performance anion chromatography), gas-liquid chromatography, and the YSI Analyzer invented by Yellow Springs Instruments (Yellow Springs Instrument, Yellow Springs, Ohio, USA). While these techniques are shown to be accurate and used for quality assurance in processing facilities, they are still destructive, time consuming, and cannot cope with in-line sorting applications. The levels of sugars in potato tubers are very

critical for estimating the viability of processing, especially chipping, and French frying. According to Stark and Love (2003), the recommended thresholds of glucose at either harvest time or during storage are 0.035% (FW) for potatoes destined to chips and 0.12% (FW) for potatoes used for French fries. Sucrose thresholds are 0.15% (FW) at harvest and 0.10% (FW) during storage for chipping tubers, whereas those values were 0.15% (FW) at harvest or during storage for tubers dedicated to French fries use. Higher levels of reducing sugars cause a dark browning color resulting from the non-enzymatic reaction, known as the Maillard reaction, between reducing sugars and the amino acid asparagine (Storey and Davies, 1992). In addition, sweetening flavor found in potato chips, and French fries is due to the increase of sucrose content as a result of storing tubers at low temperatures ($< 4^{\circ}\text{C}$) (Storey, 2007). Thus, monitoring sugars in potato tubers before, and during, storage becomes a basic quality practice in the frying industry.

Some studies of electronically assessing sugar content of potato tubers yielded relatively low values of RMSEP (Mehruabeoglu and Cote, 1997; Hartmann and Büning-Pfaue, 1998; Haase, 2011; Rady et al., 2014) that are lower than the threshold listed for processing. Other experiments conducted on whole tubers, however, either resulted in higher RMSEP values (Yaptenco et al., 2000, Rady et al., 2014) or did not include independent prediction sets (Chen et al., 2010). Such lower performance is mainly due to the skin effect that is cultivar dependent. Consequently, sorting potato tubers based on sugar content is a more challenging task than assessing sugars in ground, homogenized, or even sliced samples. Classifying tubers with respect to their sugar content reduces the variation of sugars between them and helps improve frying quality and consistency. Moreover, tubers with higher sugar content than the processing thresholds may be reconditioned by storing at elevated temperatures for 2-6 weeks (Storey, and

Davies, 1992). More research regarding the on-line sorting of potato tubers based on sugar content is still needed for enhancing the quality of both fried products, and fresh tubers.

Potato does not contain considerable amounts of proteins, 1.7-2.1 g per 100 FW, compared to eggs, fish, and dairy products. However, in countries with high potato consumption, potato significantly contributes to human diet. Moreover, the high quality of potato protein within 100 g of boiled potato supplies the portion of Recommended Daily Allowance (RDA) of 8-13% for children, 6-7% for adults (Storey, 2007; Storey and Davies, 1992). Burton (1989), also stated that potatoes yields more protein per hectare than major cereal crops. Other uses of potato protein include cattle and pig feed, as well as some other applications including treatments for weight loss, peri-anal dermatitis, thrombotic disease, and cancer (Kärenlampi and White, 2009). Therefore, estimating protein content in potatoes in a rapid way can help assess the viability of tubers for industry. The Kjeldahl procedure is the traditional method for estimating protein in food products, and it is a destructive technique requiring enough time for digestion, neutralization, and titration steps (Chang, 2010). NIR diffuse reflectance (1100-2500 nm) was successfully used by Hartmann and Büning-Bfaue (1998), to estimate protein content of potatoes with R(RMSEP) values of 0.86(0.06%), which was more accurate than results achieved by Haase (2006), using VIS/NIR reflectance (400-2500 nm) with R(RPD) values being 0.79(0.205%FW) which refers to the advantage of choosing narrower wavelength bands in the former study so that interference from other chemical compounds was reduced.

Table 2.2. Reported electronic methods to estimate carbohydrates (starch, sugars), and protein content for raw and non-processed potatoes.

Mode(spectral range)	Parameter	Tested material	R%(RMSEP, %FW)	Reference
NIR transmittance (2050-2400 nm)	Total reducing sugars	Sliced samples Russet variety Chipping variety Both	98(0.0671) (no test set) 81(0.0224) 51(0.0600)	(Mehruabeoglu and Cote, 1997)
NIR reflectance (1100-2500 nm)	Fructose Glucose Sucrose Total reducing sugars Starch Crude protein	Homogenized samples	89(0.028) (no test set) 70(0.041) 62(0.037) 82(0.061) 93(0.028) 86(0.06)	(Hartmann and Büning-Pfaue, 1998)
VIS/NIR interactance (400-1100 nm)	Glucose Fructose Sucrose Reducing sugars Total sugars	Whole tubers	83(0.087) (no test set) 95(0.101) 95(0.341) 93(0.204) 95(0.598)	(Yaptenco et al., 2000)
NMR (low field)	Starch	Raw (slices)	?(?)	(Thygesen et al., 2001)
VIS/NIR reflectance (400-2500 nm)	Starch	Mashed tubers	98(0.651)	(Haase, 2004)
NIR interactance (700-1100 nm)	Carbohydrates	Whole tubers	93(0.98)	(Chen et al., 2004)
Opto-electric system	Starch (using density)	Whole tuber	?(?)	(Hoffmann et al., 2005)
VIS/NIR reflectance (400-2500 nm)	Starch Protein Coagulable protein	Mashed tubers	95(0.740) 79(0.205) 50(0.093)	(Haase, 2006)
VIS/NIR interactance (400-1100 nm)	Glucose Fructose	Whole tubers	0.65(0.046) (no test set) 0.71(0.026)	(Chen et al., 2010)
NIR interactance (850-2500 nm)	Starch (incremental) Starch (retrospective) Reducing sugars (incremental) Reducing sugars (retrospective) Sucrose (incremental) Sucrose (retrospective) Total sugars (incremental) Total sugars (retrospective)	Ground samples	98(0.50) 98(0.47) 57(0.00483) 66(0.00389) 77(0.0106) 84(0.00969) 73(0.0156) 81(0.0135)	(Haase, 2011)
Vis/NIR interactance (446-1125)	Glucose Sucrose	Sliced samples & Whole tubers (Chipping-table use)	90-95 (0.0515-0.0786) & 88-79(0.0620-1529) 81-50(0.0439-1.0273) & 81-26(0.0436-0.2051)	(Rady et al., 2014)
Vis/NIR hyperspectral reflectance (400-1000 nm)	Glucose Sucrose		64-74 (0.0880-0.1643) & 38-52 (0.0681-0.3259) 62-57(0.0580-0.1533) & 14-43 (0.0702-0.1805)	
NIR transmittance (900-1685 nm)	Glucose	Only sliced samples	66-87 (0.0515-0.1921)&	
	Sucrose		57-63 (0.0582-0.8962)	

2.1.3 Defects and Diseases Detection of Potato Tubers

Mechanical damage and disease management are probably the most critical postharvest issues that face growers and processors. Negative consequences occur for potato products when there are inappropriate harvest and handling operations. The study of mechanical damage in potatoes was among the earliest postharvest problems addressed and presented in literature (Klapp, 1945; Hopkins, 1953; Nylund, and Hemphill, 1955; Volbracht, and Kuhnke, 1956;

Lamp, 1960; Ophuis et. al, 1958; Zahara et. al, 1961; Parke, 1963; Loow, 1964; Kunkel, and Gardner, 1965; Johnston et. al, 1968; Gray and Hughes, 1978; Hyde et. al, 1979; Balls et. al, 1982; Mohsenin, 1986; Burton, 1989; Kleinschmidt and Thronton, 1991; Baritelle et. al, 1998; Baritelle and Hyde, 1999; Thronton and Bohl, 2000; Hemmat and Taki, 2001; Rady, 2006; Rady and Soliman, 2013). Dean (1996), stated that the brown or black discoloration seen in tubers after impacts is caused by both enzymatic and non-enzymatic oxidation of phenolic substances. The enzyme called polyphenoloxidase (PPO) results in the formation of melanin pigments. According to Storey and Davies (1992), mechanical damage of potato tubers may be divided, based on the form of damage, into two groups: external or internal damage. External damage includes skin scuffing, cuts or gouges, crushing, which are apparent by inspection, and leads to direct losses when grading or preparation for consumption or processing. It also causes an increase in weight loss during storage and allows for the ingress of disease pathogens. The second type is internal damage, which includes internal shattering or cracking and black spots. In some cases, internal damages may be visible under the skin of the tuber, but in most instances it is not apparent until tubers are cut or peeled.

Defects and diseases were also some of the first postharvest problems that received much investigation into noninvasive and/or electronic techniques. These disorders usually result in change in shape, tissue color, or moisture content that can be detected using non-invasive techniques. This domain became an open field for research using rapid and/or electronic methods which led to systems already available to the industry to help sort non-desirable tubers or potato products. The reason of this early importance is the severe economic impact of such problems in either fresh or processed forms. There are many electronic-based rapid techniques applied to potato to assess defects including traditional machine vision, spectroscopic, and ultrasonic.

Application of such methods on raw potato tubers along with performance are discussed in sections 2.1.3.1.-2.1.3.3. Various electronic methods used for defect detection of potato tubers are shown in table 2.3.

2.1.3.1 Spectroscopic-based methods

Spectroscopic techniques have been used in many quality evaluation applications including detecting defects for fruits, vegetables, grains, and meat. Detection of potato defects using spectroscopic systems depends on variation of absorbance between sound and damaged tissues that is usually used to classify tubers into different categories.

Hollow heart (HH) was one of the earliest defects to be studied using noninvasive techniques possibly because it is a major internal physiological disorder that significantly affects tubers dedicated for processing. Due to the fact that HH usually develops as an irregular cavity in the pith area (Watts and Russel, 1985), the transmittance mode was probably the appropriate technique for detecting such defect. Several factors, however, resulted in somewhat low classification rates of HH (83-98% for Birth 1960; and 83% for Kang et al., 2008) using spectroscopic methods as noted in table 2.3. The most influencing factor for such results is the similarity of absorption characteristics between skin and damaged tissue (Birth, 1960). Consequently, some small tubers were classified as false-positive as a result of the fact that the proportion of path length through the skin with respect to the total path length is higher for small tubers than larger tubers. Other internal defects (black spot) followed the same results as for HH.

Some defects have internal breakdown of the tissue extending to the surface (bacterial soft rot, dry rot, late blight, gangrene) and were also classified using spectroscopic methods resulting in comparable performance to that of HH, and black spot (Muir et al., 1982).

In general, internal defects were much more successfully detected using different spectroscopic systems than external defects due to specular reflectance and interference from tuber skin in the latter type. Moreover, external defects are usually not completely distributed over the tuber surface which requires scanning of the whole surface to obtain accurate description of tuber status and consequently a high classification rate.

2.1.3.2 Imaging-based methods

Applications of computer vision systems on detecting physiological disorders, mechanical damage, and other internal or external defects of potato tubers were studied to evaluate the potential of using such techniques for sorting tubers dedicated for either fresh use or processing. Hollow heart, bruises, greening, scab infection, and blemishes are probably the most frequent imperfections that received consideration of imaging-based methods as shown in table 2.3.

As a result of its efficient use in medical diagnostics, x-ray imaging systems were dominating computer vision research studies in the agriculture domain since the 1930's. X-ray is a short-wave electromagnetic spectrum (0.002-100 nm) that interacts with specimen tissue and the intensity of detected signals mainly depends on incident intensity, absorption coefficient, product density, and sample thickness (Butz et al., 2005; Abbott, 1999). Studying the detection of hollow heart in potatoes was the first application of x-ray in quality measurements for perishable produce (Abbott, 1999; Nylund, and Lutz, 1950; Harvey, 1937). Experiments conducted by Nylund, and Lutz (1950), Finney and Norris (1973 and 1978), resulted in classification rates of 84.1, 100, and 100% respectively for defected tubers.

Some challenges still restrict the application of x-ray imaging systems in the domain of food products. Such restraints include the limitation of detection to density-changing tissues and not chemical composition or mechanical damage forms, the high cost of x-ray inspection systems, and low operational speed (Mathanker et al., 2013; Butz et al., 2005; Chen et al., 2002).

Rapid development of imaging hardware and computers resulted in the application of color cameras on tracking quality attributes of food products. Images resulting from color cameras show useful information about both internal and external status of samples. Obtained information includes color, shape, textures, disease, and defects. With the decreased cost, and increasing computing speeds, image analysis was made possible for building commercial grading systems for fruits, and vegetables (Chen and Sun, 1991).

Several studies were conducted to investigate the potential use of color cameras, along with other imaging systems for defect detection of potatoes. Surface defects (skin cutting, shatter bruise, common scab, greening, cracks, etc.) were successfully evaluated for whole tubers using color cameras with classification rates higher than 95% (Hasankhani et al., 2012; Samanta et al., 2012). Other internal or sub-surface defects were also studied using RGB, and multispectral cameras, or hyperspectral imaging systems. Results of classifying common scab defected tubers and healthy tubers using NIR hyperspectral imaging by Dacal-Nieto et al. (2012), showed promising performance with classification rates of healthy and defected tubers of 94.0%, and 98.6% respectively. It is worth stating that hyperspectral imaging systems are not suitable for on-line sorting purposes because the relatively long acquisition time needed to acquire each image. They can be effectively used to provide the most influencing wavelengths associated with the high classification rates, and those wavelengths can be utilized by multispectral imaging systems (Chen et al., 2002).

Several techniques were noted above for effectively monitoring different external and internal disorders in potato tubers with various degrees of efficiency. Imaging systems were noted to present the best performance of tracking the presence of defects and damage compared to other systems, especially for internal defects. Although acquiring and analyzing spectroscopic signals is less time consuming than for imaging systems, the use of spectroscopic systems for detecting internal defects, that are not visible by human labor, did not yield acceptable performance for the industry.

2.1.3.3 Sonic-based methods

Ultrasound technology (UT) is known for its successful use in medical diagnosis, and manufacturing applications. UT usually works under either of two modes; the pulse-echo mode which is simply a reflectance mode in which one transducer is used for emitting and receiving the reflected signals. In the second mode, known as the through-transmission mode, one transducer works as a transmitter and the second one as a receiver. Evaluation of tested material using UT comes from both attenuated signals and the propagation speed as both parameters vary with the change of tissue nature or the presence of defects (Mizrach, 2012; Mizrach, 2008). Unlike solids, liquids, and human tissues, fruits and vegetables are very attenuating materials due to their scattering effect when applying the frequencies used for medical and industrial applications (0.5-30MHz) (Mizrach, 2008). More studies by Sarkar and Wolfe (1983), reported that lower frequencies (100-500 KHz) and higher acoustic power might be more effective for quality applications of fruits, and vegetables.

As presented in table 2.3, ultrasound technology was generally applied for potatoes on detection of hollow heart as this physiological disorder tends to have distinguished wave attenuation characteristics compared with healthy tissue and generally, defective tubers had less

signal amplitude and intensity than the healthy tubers. Most studies were conducted in the frequency range of 50-200 KHz. Success in hollow heart detection was demonstrated with a classification rate as high as 100% (Ha et al., 1991; Cheng and Haugh, 1994). Such results showed the advantage of using ultrasound techniques for detecting hollow heart and possibly other diseases and damage in potato tubers. Limitation of tuber defects that can be effectively tracked using UT, however, restricted the application of UT to hollow heart only which is not economically valuable with the many diseases and disorders infecting potatoes in the postharvest stage as mentioned earlier.

Table 2.3. Reported spectroscopic, imaging, and sonic methods to detect defects, and diseases for potato raw and non-processed potatoes.

Mode(spectral range)	Defect/disease	Classification rate (%)	Reference
Transmittance (540-910nm)	Hollow heart Greening Decay Black spot	83-98 50 50 88	(Birth, 1960)
VIS/NIR diffuse reflectance (590-2030 nm)	Defected tubers Sound tubers	79 82	(Porteous et al., 1981)
VIS/NIR diffuse reflectance (570-870 nm)	Gangrene (control) Gangrene (diseased) Dry rot (control) Dry rot (diseased)	98 77 93 72	(Muir et al., 1982)
VIS/NIR diffuse reflectance	Surface & subsurface defects	?	(Muir et al., 1999)
UV to NIR (250-1750 nm) reflectance UV VIS NIR	Surface bruise Unpeeled Peeled Unpeeled Peeled Unpeeled Peeled	 45.5 79.5 55.1 57.1 65.9 55.8	(Evans and Muir, 1999)
VIS/NIR transmittance (530 – 1100 nm)	Hollow heart	83	(Kang et al., 2008)
VIS/NIR time resolved reflectance (540-900 nm)	Internal brown spot	81	(Vanoli et al., 2012)
X-ray	Hollow heart	?	(Harvey, 1937)
X-ray	Hollow heart	84.1	(Nylund and Lutz, 1950)
X-ray	Hollow heart	100	(Finney and Norris, 1973)
X-ray	Hollow heart	100	(Finney and Norris, 1978)
CCD color camera	Greening Shatter bruise	74.0 76.7	(Deck et al., 1995)
CCD color camera	Greening	90.0	(Tao et al., 1995a)
Multispectral camera (400-2000 nm)	Surface & subsurface defects	?	(Muir et al., 1999)
Color camera	Colored bruises & greening	?	(Marique et al., 2005)
CCD color camera	Good potato Potato with defects	100 100	(Jin et al., 2009)
CCD color camera	Blemishes White cultivar Red cultivar	 89.6 89.5	(Barnes et al., 2010)
NIR Hyperspectral 900-1700 nm	Healthy tubers Common scab	94.0 98.6	(Dacal-Nieto et al., 2011)
CCD camera	Greening	94.7	(Ebrahimi et al., 2011)
CCD color camera	Healthy Crack Greening Fetidness Skin cutting Other defects	100 100 100 86.0 100 100	(Hasankhani et al., 2012)
CCD color camera	Defected tubers	95.0	(Razmjoooy et al., 2012a)
RGB camera	Scab disease	97.5	(Samanta et al., 2012)
Ultrasound attenuation at 175 KHz	Hollow heart	?	(Watts, and Russell, 1985)
Ultrasound attenuation at 50, 100 KHz	Data collection	?	(Mizrah, 1989)
Ultrasound attenuation at 50 KHz- 1 MHz	Hollow heart	100	(Ha et al., 1991)
Ultrasound attenuation at	Hollow heart	?	(Mizrach, et al., 1992).
Ultrasound attenuation at 250 KHz	Hollow heart	100	(Cheng and Haugh, 1994)
Ultrasound attenuation at 50,100,150 KHz	Hollow heart	98	(Jivanuwong, 1998)
Acoustic impact	Hollow heart	98	(Elbatawi, 2008)
Ultrasound attenuation at 2, 32.8, 40 and 50 KHz	Mechanical damage (pressure, and impact)	83-95	(Eseghbeygi et al., 2011)

2.1.4 Systems for Non Destructive Sorting of Raw Potato Tubers

Elimination of tubers with surface defects, physiological disorders, and/or internal drawback that don't meet quality requirements is a necessary process during preparing potatoes for fresh market or processing. US Standards for Grades of Potatoes restrict potato growers with limits for defects, size, weight, maturity, and shape uniformity with tolerances either in the shipping or destination points with different grades including US. No.1, U.S. Commercial, and U.S. No. 2 (National Agricultural Statistics Service, USDA, 2012). Although there are regulation tolerances, proper considerations should be taken at sorting stations and packing houses to ensure higher product marketability and more benefits for producers.

The use of human labor for sorting and grading of agricultural products is the traditional technique especially in developing countries as the labor is much cheaper than in developed countries. Some disadvantages, however, are usually related to manual sorting including low sorting rate, inability to sort internally defected samples, degradation of performance with time, and the high cost and problems associated with immigrant workers in developed countries. Such drawbacks resulted in extensive research for developing techniques for detecting defects and physiological disorders in a noninvasive manner, as shown in section 2.1.3, and quantitatively and qualitatively improving the output of sorting stations. Sorting potato tubers, as well as other perishable products, is more complicated than the prediction of a single or multiple constituents or the detection of one or more defects. In designing any sorting system, one should consider not only important factors stated in section 2.1.3, but also other foreign materials that need to be discarded such as vine, stones, soil, etc. Moreover, a sorting process includes eliminating samples that don't match size, appearance, and shape standards. Also, productivity of the sorting system (ton/hr) is a crucial factor to estimate its practicality, and marketability.

In the case of potatoes, several operations are conducted on the harvested crop including removal of any remaining soil, clods, haulm, and stones, grading of tubers into several sub-grades, and removing any tubers that do not meet requirements for local market (Pringle et al., 2009). Table 2.4 shows the different reported studies for sorting potato tubers using several techniques (spectroscopic, imaging, ultrasonic, vibrational response characteristics).

Spectroscopic methods are known to be rapid techniques for monitoring quality attributes for potatoes as shown in sections 2.1.1, 2.1.2, and 2.1.3. With the relatively low integration time, i.e. high acquisition speed, for the developed diode array-based NIR spectrometers, in addition to the powerful multivariate regression methods, i.e. PLSR, it was feasible to build online spectroscopic sorting systems (Nicolai et al., 2007). The most crucial factor affecting the performance of such a sorting system is the prediction model robustness that includes the ability to precisely predict quality attributes for samples that were not used to build the calibration model. The calibration models should be based on large datasets obtained from different destinations, growing conditions, and operational conditions (Nicolai et al., 2007).

Table 2.4. Reported electronic methods for sorting potato raw potato tubers using different quality attributes.

Method/ Mode	Sorting criteria	R(RMSEP) or classification rate (%)	Performance	Reference
VIS/NIR diffuse reflectance (600-1300 & 1500-2200 nm)	Reflectance characteristics	100(for potatoes)	292320 objects/hr	(Story and Raghavan, 1971)
NIR diffuse reflectance	Reflectance characteristics	?(?)	?	(Story, 1973)
VIS/NIR transmittance (530 – 1100 nm)	Dry matter Specific gravity	80.0(0.67 %) 83.0(0.005)	?	(Kang et al., 2008)
NIR diffuse reflectance (1100-2500 nm)	Dry matter Starch Coagulating protein	97.0(0.47%) 92.0(0.63%) 92.0(0.06%)	?	(Brunt and Drost, 2010)
X ray absorption and scatter	Absorption and scattering	?	?	(Slight, 1966)
TV camera	Size and shape	100.0	40 tubers/sec	(Marchant et al., 1990)
CCD color camera	Color Green tubers Good tubers	 90.0 100.0	?	(Tao et al., 1995a)
CCD color camera	Shape	89.2	?	(Tao et al., 1995b)
CCD video camera Moving tubers Stationary tubers	Shape and size	 88.0 98.0	3 tubers/min	(Heinemann et al., 1996)
CCD video camera	Weight Shape Size Color (greening) Overall	91.2 89.1 85.5 78.0 87.0	50 potato images/sec	(Zhou et al., 1998)
CCD color camera	Color (greening) Shape Surface crack Rhizoctonia	88.1 99.2 100.0 100.0	50 tubers/sec 12 ton/hr	(Noordam et al., 2000)
RGB camera Hyperspectral imaging (321-1044 nm)	Shape Clods (with wet, dry tubers) Wet tubers, dry tubers Clods (with wet, dry tubers) Wet tubers, dry tubers	 94.4, 75.4 91.2, 71.4 99.8, 97.4 100, 76.8	?	(Al-Mallahi et al., 2008a)
UV CCD camera (300-380 nm)	Color Clods Tubers	 71.2-100 94.5-100	?	(Al-Mallahi et al., 2008b)
CCD color camera	Defects & color Good Rotten Green	 86.6 88.7 86.2	?	(Dacal-Nieto et al., 2009)
Firewire camera	Shape	93.8	?	(Rios-Cabrera et al., 2009)
UV CCD camera (300-420 nm)	Color Clods Tubers	 100 100	?	(Al-Mallahi et al., 2010a)
UV CCD camera (300-420nm)	Color Clods Tubers	 98.3 98.8	?	(Al-Mallahi et al., 2010b)
CCD camera	Size (minor and major axis)	?	?	(Chenglong et al., 2011)
CD color camera	Shape Regular tubers Misshapen tubers	 98.8 75	?	(ElMasry et al., 2012)
CCD color camera	Shape Accepted tubers Rejected tubers	 91.8-100 100	?	(Hasankhani and Navid, 2012)
CCD color camera	Shape Defects	96.9 95.0	?	(Razmjoooy et al., 2012)
Vibrational response characteristics	Vibrational response (clods)	?	?	(Miller and Stephenson, 1971)
Vibrational response characteristics	Vibrational response Clods (static tests) Clods (moving objects)	 100 90-100	4-5 Pocket/sec	(Stephenson et al., 1979)
Impact acoustic signals (up to 100 KHz) Off-line On-line	Impact acoustic response Tubers Clods Tubers Clods	 97.3 97.6 97.2 97.5	20 ton/hour	(Hosainpour et al., 2010)

Review of literature showed that sorting potato tubers from stones, clods, and other foreign materials was not effectively studied using spectroscopic systems except in two references in which VIS/NIR, and NIR diffuse reflectance characteristics of potato tubers, stones, and clods were studied by Story (1973), and Story and Raghavan (1971), respectively. Both investigated the difference of diffuse reflectance properties between tubers and other foreign materials. Although the classification rate in the former study for potato tubers was 100%, the authors stated several problems that could reduce system performance including the detector balance and the heating transistor. Moreover, it is important to state that results of such study were not confirmed by further research or applied on different cultivars. The possible reasons for no further advancement in this area includes the deficiency of spectroscopic systems in grading tubers based on size, and shape. On the other side, the rapid improvement of imaging hardware resulted in fast and accurate identification of size, and shape of different objects.

As shown in sections 2.1.1, and 2.1.2, spectroscopic systems have been studied to estimate different quality attributes of potatoes. Studies were conducted in an attempt toward developing an on-line sorting system for potato tubers based on quality parameters that are associated with processing. Dry matter (DM) and specific gravity (SG) were studied as sorting criteria by Kang et al. (2008), using VIS/NIR transmittance (650-1000 nm) with R(RMSEP) values for the prediction set of 0.83(0.0050), and 0.80(0.0067) for SG, and DM respectively. NIR diffuse reflectance (1100-2500 nm) was also used by Brunt and Drost (2010), for obtaining prediction models of dry mater, starch, and coagulation protein for potatoes in an off-line mode in an attempt to build a sorting system. Values of R(RMSEP) of prediction models were 0.97(0.47%), 0.92(0.63%), and 0.92(0.06%) for dry mater, starch, and coagulation protein respectively.

Machine vision applications in grading and sorting of vegetables and fruits based on size, color, weight, and defects were stated in the literature (Mahendran et al., 2012; Vibhute and Bodhe, 2012; Chen et al., 2002; Abbott, 1999; Chen and Sun, 1991). In the case of potatoes, there is a considerable variation of size, shape, and color of the cultivars available in the local market which adds difficulty for building a robust, yet universal sorting machine that can tolerate such changes. Several studies were conducted to build systems able to sort tubers based on different quality attributes using computer vision techniques as illustrated in table 2.4.

Experiments on sorting potato tubers using imaging methods started as early as 1966 by Slight in which x-ray scattering and absorption characteristics were studied under low energy (40 KeV) to differentiate between tubers and rocks. While absorption coefficient values for potato tissue were less than those for other materials which gave a possibility for sorting potatoes from rocks, no further studies were conducted to enhance the results with a possible reason being the rapid development of imaging hardware, especially CCD-based cameras. Thereafter, studying the potential use of imaging systems in potato sorting and grading were extensively accelerated. Greening, as an external defect, was successfully detected based on tuber surface color by Tao et al., (1995); Zhou et al., (1998); Noordam et al., (2000); and Dacal- Nieto et al., (2009), with classification rate of defected tubers of 90%, 78.0%, 88.1%, and 86.2% respectively. Sorting and grading tubers could be a difficult mission with the singulation problem as a result of the possible interference between different touching objects (Al- Mallahi et al., 2010a; Marchant et al., 1990). It was possible, however, to build grading systems for tubers based on size by developing several separating techniques, applied on the captured images, such as the blob splitting algorithm (Marchant et al., 1990), the 8-neighbor labeling algorithm (Al- Mallahi et al., 2010a), or based on intensity threshold (Dacal-Nieto et al., 2009). Consequently, grading tubers

into several standard grades, and eliminating misshapen tubers, was successfully conducted by Marchant et al., 1990; Noordam et al., 2000, and El Masry et al., 2012, with classification rates of 100%, 99.23%, and 98.8% respectively. Separating clods, stones and other foreign materials is another application studied by imaging systems. Achieving high detection rate of clods was possible using hyperspectral imaging (321-1044 nm) by Al- Mallahi et al. (2008a), in which the rate was 99.8%, 97.4% in the wet, and dry conditions, or UV camera (300-420 nm) by Al- Mallahi et al. (2010a), with the rate of 100%. The application of imaging systems for sorting potatoes based on external defects was also investigated by Noordam et al., (2000); Dacal-Nieto et al., (2009); and Razmjoo et al., (2012), with classification rates of the defected tubers being 88.7%-100%.

Applying vibrational response characteristics on sorting potato tubers from clods, stones, or defects was initially studied by Miller and Stephenson (1971), by exciting a mixture of potato tubers, clods, and stones using either electromechanical or sonic techniques in the range of 20 to 2 KHz. Results showed differences of vibrational response between the three objects. Additional study by Stephenson et al. (1979), assessed resonant frequencies of several perishable products among which potatoes show a frequency band of 400-600 Hz to yield the best detection of clods and stones (100% in static mode, and 90-100% in moving mode).

2.2 Applications of Non Destructive and/or Rapid Methods on Quality Evaluation for Potato Products

Processed potato products are more consumed in developed countries compared to table use in developing countries. However, noticeable changes are occurring in the developing countries toward consuming processed products. With advances in frozen French fry manufacturing facilities since the 1950s and the increase in fast food chains, processed potatoes

contributed to 64% of the total US potato use in the 2000s with 39 pounds per capita compared to 19 Kg per capita for fresh tubers (Economic Research Service (ERS), 2012). Moreover, according to a NASS 2013 report, US potato utilization included 61.2% of the total 2012 crop production directed to processing, 25.6% sold as fresh tubers, and 5.8% used as seed. Among the processed tubers, frozen French fries contribute to 51.1%, and 20% for chips and shoestrings. Moreover, in 2012, French fries, and chips contributed to 74% (> \$1 billion) of the total US potato exports to the global market (2013 potato statistical yearbook). Thus, in this section the application of rapid and/or nondestructive methods on assessing quality attributes of processed potato will be limited to these two products.

During the frying process, a significantly different microstructure, compared to the raw tuber, is derived. Formed after the frying process, the surface of a chip or French fry becomes dry, crispy and oily. However, the inner part is moist and cooked with less oil content than the external surface. Moreover, the oil content in the potato chip and French fry is around 38%, and the moisture content is 1.8 and 15% respectively (Pedreschi, 2009). However, in the case of raw tubers, the moisture, and fat contents are 77%, and 0.5% respectively (Kadam et al., 1991). Also, processed products are in general more uniform and controlled in shape and dimensions compared to raw tubers. Thus, it can be concluded that quality assurance strategies for French fries and chips are significantly different from raw tubers.

Frozen French fries and chip marketability is extremely affected by the appearance which is the first factor influencing customer evaluation for the final product. Thus, quality assurance requires French fries and chip color to follow standards established by USDA, and other governmental and/or industry-related organizations. Frying color is affected by many factors including cultivar, maturity, stress during growth, storage period temperatures, handling

practices, fertilization application, slice thickness, frying parameters, and moisture content of the final product (Gould, 1995b).

The capability of spectroscopic systems to assess and identify many chemical and nutritional compounds using their specific absorption signature at definite wavelengths has resulted in extensive research for quality monitoring of French fries and chips. Table 2.5 shows the reported spectroscopic methods for studying several quality attributes of French fries and chips.

To assess French fry color, visible diffuse reflectance (400-700 nm) was utilized by Panigrahi et al. (1996), and extracted features included color and reflectance properties. By applying linear discriminant analysis (LDA), classification accuracy was as high as 86, 86, and 100% for dark, normal, and light groups respectively.

A potato chip is very thin (1.27-1.78 mm) and several quality attributes are important to monitor throughout production. Dry matter is an effective factor in frying oil consumption and dehydration during frying (Storey and Davis, 1992). Fat content in fried products is an important concern for consumers as healthy food is a major target in the current human diet around the world (Pedreschi, 2009). Although the Maillard reaction is known for formation of browning color during the frying process, another component is also formed, which is acrylamide, discovered by the Swedish National Food Authority in 2002 (Mottram et al., 2002; Stadler et al., 2002). The acrylamide single unit (monomer) is toxic to the nervous system, a carcinogen in laboratory animals and a possible carcinogen in humans. French fries, and chips contain fat concentrations of 424 µg/kg, and 1739 µg/kg and these are considered relatively high ratios (Pedreschi, 2009). Spectroscopic systems are known for their efficacy to qualitatively and quantitatively monitor chemical components in food products. Consequently, several research

studies were conducted to assess these quality attributes. Prediction of acrylamide in potato chips was successfully conducted using VIS/NIR diffuse reflectance (400-2500 nm) by Segtnan et al. (2006). Results showed high performance for prediction models with R (RMSEP) of 0.95(246.8 µg/kg). Such results were better than those obtained by Pedreschi (2010b), (460-740 nm & 760-1040 nm) which is possibly due to the extended range of wavelengths used in the former study. Fat and moisture contents in potato chips were also evaluated by Shiroma and Rodriguez (2007), using NIR and MIR spectroscopy (1052-2000 nm and 2500-13333 nm) and best results showed R(RMSEP) values of 0.97(0.3%), and 0.96(1.29%) respectively. Pedreschi (2010b), was able to assess fat and dry matter contents with prediction models having R(RMSEP) values of 0.99(0.99%) and 0.97(0.84%) respectively. Evaluation of quality attributes of French fries and potato chips using spectroscopic systems, as mentioned above, presented a potential for building sorting systems, or handheld tools for rapid assessment of both products after frying and before packing which increases the final product grade and also presents healthier food for consumers.

Imaging systems are extensively used for color-based sorting and defect detection for multiple food products for their efficiency in detecting color differences using inexpensive cameras. Moreover, considering color as the most apparent, yet crucial quality aspect for both French fries and chips, much attention was carried out toward studying the potential establishment of cost-effective sorting imaging systems for these two potato products as shown in table 2.5.

Table 2.5. Reported spectroscopic and imaging methods for assessing quality attributes for frozen French fries and potato chips.

Method/ Mode	Product	R(RMSEP) or classification rate	Reference
Visible diffuse reflectance (400-700 nm)	French fries/ reflectance properties Overall Light Normal Dark	91% 100% 86% 86%	(Panigrahi et al., 1996)
VIS/NIR diffuse reflectance (400-2500 nm)	Potato chips acrylamide content	95(246.8 µg/kg)	(Segtnan et al., 2006)
NIR(1052-2000 nm) MID-IR (2500-13333 nm)	Potato chips Fat Moisture content Fat Moisture content	97(0.3%) 96(1.29 %) 97(0.3%) 96(1.65 %)	(Shiroma and Rodriguez-Saona, 2007)
VIS/NIR Inductance (460-740 nm & 760-1040 nm)	Potato chips Fat Dry matter Acrylamide	99(0.99%) 97(0.84%) 83(266 µg/kg)	(Pedreschi et al., 2010b)
Video camera	Potato chips (color measurement)	99	(Coles et al., 1993)
Video camera	Potato chips (color measurement)	94	(Scanlon et al., 1994)
Video camera	Potato chips (color defects)	?	(Segnini et al., 1999)
Digital color camera	Potato chips (color defects)	90	(Marique et al., 2003)
Multispectral (400-900 nm) CCD color camera	French fries (defects)	87.90-99.25 69.3-93.9	(Noordam et al., 2004)
Digital color camera	Potato chips (color)	90-100	(Pedreschi et al., 2004)
Digital color camera	French fries Internal hollowness Normal Total	100 100 100	(Yin and Panigrahi, 2004)
Digital color camera	Potato chips (color & frying temperatures)	?	(Pedreschi et al., 2006)
Digital color camera	Potato chips (color and texture)	90	(Mendoza et al., 2007)
Flatbed scanner	Potato chips (color defects)	98	(Romani et al., 2009)
Digital color camera	Potato chips (color) Smooth chips Chips with ruffles	97 82	(Pedreschi et al., 2010a)

Most studies of investigating chip color using imaging techniques were conducted using either digital or video cameras as they are relatively inexpensive, and at the same time can efficiently detect color differences. Correlation between color features, especially in the L*a*b space that is more human-related and less dependent on illumination (Segnini et al., 1999), and measured color of chips were extensively conducted to evaluate external quality and estimate the presence of any undesirable dark color spots. Romani et al. (2009); Marique et al. (2003); and Scanlon et al. (1994), applied such techniques with R values of 0.98, 0.90, and 0.94. Surface shape of chips was proven to reduce such correlation as shown by Pedreschi (2010a), in which the R values were 0.97 and 0.82 for smooth and undulated chips. Mendoza et al. (2007), found that texture-based features (energy, entropy, contrast, and homogeneity) yielded better

classification rates (90%) than using color-based features. A combination of color and texture features was selected using Fisher linear discriminant functions and resulted in high accuracy (90-100%) for classifying chips into different classes based on frying parameters.

The application of imaging systems on French fries was restricted to the area of defect detection. Multispectral imaging (MI) (400-900 nm) showed higher classification rate (87.90-99.25%) than RGB color (69.3- 93.3%) for assessing several defects (damage, greening, external rot, and browning). The possible reason for such a trend was the ability for MI to identify some defects not shown in RGB images (i.e. greening). Texture features were also utilized for detecting hollowness in French fries using an RGB camera (Yin and Panigrahi, 2004). Ideal classification (100%) was obtained for normal, and defected strings using features obtained from gray level images along with a co-occurrence algorithm for feature calculation. However, computation time, as a crucial factor to assess the applicability of such a method for developing on-line sorting system, was not addressed in this study and needed further investigation.

2.3 Commercial Sorting Systems for Potato Tubers, French Fries and Chips

During the last three decades, nondestructive systems for sorting perishable products were successfully transferred from research labs into fields, packing houses and processing plants. Potatoes were one of the most applicable commodities to receive attention to apply rapid and noninvasive technology to discard internally or externally defected, misshapen and non-suitable sized tubers. Also potato chips and French fries were classified using systems available commercially. Table 2.6 shows commercial sorting systems available in the market for whole potato tubers, French fries and chips.

Table 2. 6. Commercial sorting and quality monitoring systems for raw potato tubers, French fries and chips.

Company/model	Effective sorting base / and sorted material	Discarded materials	Notes
Odenberg/ FPS 1200,1400,1800	Multispectral NIR color cameras/ unwashed potatoes (red, brown, and white skin)	Soil clods, stones, foreign materials, and rotten potatoes	15-70 ton/hr
Key Technology/Optyx® WPS	Laser and high performance color cameras/unwashed whole tubers	Foreign materials, and rework potatoes	45 ton/hr
Key Technology/Optyx®	Multiple laser and cameras (VIS/IR, UV, or tri chromatic detection bands) configuration with LED, HID, or UV lighting /Whole tubers, chips, French fries, diced potato, wedged, and sliced	Color, shape, texture, and defects	6-12 ton/hr
Compac/InVision 9000 blemish	Vision system/ washed red and white tubers)	External defects (marks, stains, insect damage, bruises, cuts, punctures), size, weight, color, and shape	3.5 ton/hr/lane
Taste Tech/T1	NIR diffuse reflectance, NIR transmittance/whole tubers, and chips	Internal defects (hollow heart, black spot, internal browning, sugar, and Zebra chips concentration)	
VISAR/VACS 20.0 (adopted from carrot grading system)	Color imagery system/whole tubers	Non uniform shapes, greening, surface defects (rot, cracks, dark spots)	
Herbert Engineering/DDS 1200S Auto sort	Three CCD color cameras/whole tubers	Shape, size, defects: greening, rot, spots, cuts, skin discoloration.	Up to 6 ton/hr (20-40 mm size)
Oculus/ 1300 and 2000	Infrared and digital cameras/whole tubers	Defects: bruises, greening, cracks, black spots, Rhizoctonia, sliver scurf, rot, skin spot and foreign materials	25 and 37 ton/hr
Odenberg/Titan II	Infrared and digital cameras/whole tubers	Defects: bruises, greening, blemishes. Color, and size and foreign materials	11-50 ton/hr
Odenberg/Halo	LED, CCD camera, and NIR sensors/whole tubers (skin on or peeled)	Defects: bruises, greening, blemishes. Color, size and foreign materials	14-70 ton/hr
Odenberg/ Sentinel	Color cameras and NIR sensors/whole tubers	Shape, size, surface discoloration, defects and foreign materials	30-50 ton/hr
Best/ Genius Optical sorter	Cameras (monochromatic, color), laser (fluorescence, SWIR, or Detox) with LED, UV or IR lighting / French fry, and chip	Detects defects based on color, shape, structure, fluorescence and biological characteristics	
Best/POM/DYN size analyzer	Color camera with LED illumination source / French fry	Detect defects based on shape, size, and color	Up to 60 Kg/hr

General components of electronic sorters are: feeding unit which is usually a movable conveyor passing objects into the examining unit that contains the vision or optical system located in a closed box, the separating unit which is responsible for classifying different objects into the required classes, and the software that manages the sorting process. As objects move, they are scanned, often multiple times. Based on the adjusted thresholds, a decision is taken to discard foreign materials as well as samples that do not meet the set configurations. Finally, rejected objects are separated from the desired samples using either pneumatic-based or electro-mechanical fingers. Many of the sorting systems combine color cameras with spectroscopic sensors with proper use of lighting source. These systems help detect external defects (greening,

cuts, bruises, surface physiological disorders), misshapen tubers, foreign materials (clods, stones, soil pieces, vines, etc.) as well as internal defects (hollow heart, black spot, brown spot, etc.). Cameras are usually positioned at different locations around the moving belt, thus, when the tubers fall or are projected in free air while passing detectors a complete visualization of each object can be obtained. The combination of sensors yields a decision about the object status whether to be rejected or accepted. Spectroscopic sensors used in raw potato tuber or processed product sorting systems are either NIR or laser with a note that most defects of French fries and chips are external. Laser light sources are known for their concentrated, purity, high intensity, coherent, and narrow bandwidth. Laser (light amplification by simulated emission of radiation) is also distinguished for its ability to detect extensively small concentrations of species in the atmosphere. Consequently, many applications of laser were already in place in various medical, communication, and industrial areas (Friedman and Miller, 2003; Skoog et al., 2007). Agricultural applications of laser started in surveying and currently detection of defects in fresh produce is possible using sorting systems that integrate laser with other spectroscopic and/or image systems.

Although there has been success in manufacturing commercial sorting systems for potato tubers and products, constituent-based sorting is still a moving research area. Processing potato tubers for chipping or French frying requires continuous monitoring of sugars to assure high quality final product. Consequently, sorting based on defects is not enough to maintain such quality and a need for robust internal composition separation continues to increase.

2.4 Future Research

Monitoring processing-related constituents of potato tubers is an important task for storage managers to accurately track concentration of such compounds and parameters. Although

other accurate methods exists for measuring sugar content, i.e. HPLC or GC-MS, there is need for a handheld device that can be calibrated for measuring chemical (glucose and sucrose), and physical (dry matter, specific gravity) quality attributes. Recent research studies regarding the use spectroscopic systems showed that such a device can work with whole tubers or sliced samples, thus, it requires low preparation time and implements an integration time of less than 100 ms. Consequently, rapid measurement is feasible especially if the device is calibrated to work with multiple cultivars and different shapes, and by using large data sets and appropriate preprocessing techniques (Nicolai et al., 2007). The success of inventing a portable device would also benefit potato growers to estimate the suitable time for harvest based on monitoring different quality attributes, such as dry matter and sugar content, which are significantly affected by the pre-harvest practices as well as storage conditions.

Based on information available in literature and in the market and industry, online sorting of potatoes was mostly conducted based on eliminating foreign materials, misshapen, and defects tubers. Sorting tubers with respect to chemical constituents, and more specifically sugars, is not adequately studied. The importance of sorting tubers based on sugar content raises when potatoes from different destinations and growing conditions are stored together. Negative consequences occur, as mentioned in 2.1.2, with higher sugar concentrations when fried, thus requiring more attention in eliminating tubers with unacceptable sugar content so that they can potentially be reconditioned. Although constituent-based sorting has been used for several fruits if the target is to obtain much sweeter packed samples, however resulting in higher packing costs; in the case of potatoes, sorting based on sugar content is not only important for enhancing flavor and color quality of fries products, but it also helps provide healthier food for consumers to avoid any

subsequent problems due to acrylamide content. Thus, in the future, it may be feasible to see bags of raw potatoes in the local market marked with a “sugar-based sorted” label.

Hyperspectral imaging (HI) is also an advancing technology that has already been applied in remote sensing and precision agriculture and currently there is considerable research to apply this technology in food quality assurance. There have been few studies on assessing quality attributes of potatoes, and its fried products using HI as clarified in sections 2.1.2, 2.1.3, and 2.1.4. While this method still cannot compete with other traditional vision or spectroscopic systems in speed, it has the advantages of combined spectroscopy and imaging techniques which can work in sorting of internal or external defects as well as chemical composition. Hyperspectral systems can, however, be used as a tool to estimate optimal spectral bands for sorting based on specific criteria which can be applied in an on-line way using multispectral systems that can provide appropriate speed for commercial use (Chen et al., 2002).

There are other nondestructive techniques having feasible potential for use in quality assurance of potatoes. Magnetic resonance imaging (MRI) depends on the response of some nuclei, especially hydrogen in the case of agricultural crops, to an applied pulse of radiofrequency (RF). Images created by MRI can provide effective detection for defects resulting from watercore, bruising, or core breakdown (Abbott, 1999). Thus, MRI has been successfully applied for defect detection in apple and peaches (Barreiro et al., 1998), tomato (Milczarek et al., 2009), and pears (Hernández-Sánchez et al., 2007). X-ray computed tomography (CT), commonly used in medical applications, apply the traditional x-ray technique but over several non-parallel paths through the objects and yields a 3D projection that results in slices of such projection (Abbott, 1999). X-ray CT was used for assessing tomato maturity (Brecht et al., 1991), defects in chestnuts (Donis-González et al., 2012); and several agricultural commodities

(Donis-González et al., 2014) which gives a possible chance for external and internal defect detection in potatoes. Moreover, the application of x-ray CT imaging in French fries, and chips microstructure is feasible as the same technique was successfully applied on studying the woolly breakdown in nectarines (Sonego et al., 1995). Despite the relatively higher cost for establishing sorting systems based on NMR or x-ray CT systems, compared to spectroscopic and other vision systems, building small scale systems for grading, sorting, and quality assurance can possibly be achievable. Such possibility is due to the proven efficacy for NMR or x-ray CT systems in quality evaluation for fruits and vegetables.

2.5 Summary

Demand for processed and fast food has been showing significant increase in both developed and developing countries over the last three decades. Potato is a major crop in the food industry with various consumption forms compared with grain crops, fruits and vegetables. Potato, as other perishable commodities, is always susceptible to external and or internal damage during pre-harvest, harvesting, handling, and storage operations. Non-destructive, and/or rapid techniques of detecting defects and monitoring quality for raw tubers and processed potato products were studied first using machine vision systems with x-ray and later using spectroscopic systems.

With the advancement in vision and electronic hardware accuracy, resolution, robustness, reproducibility, and the tremendous jump in computing speed in the last decade, it has been possible to build commercial sorting systems efficient enough to eliminate external defects (physiological and mechanical) and sort tubers based on size and shape so clods, stones, and remaining vines are discarded and different tuber grades could be obtained.

The commercial use of spectroscopic systems either individually or integrated with vision systems resulted in the ability to detect and eliminate internally damaged tubers in different conditions (peeled or with skin on) which is a very important factor in assuring the quality of chips and French fries. Monitoring potato tubers for processing (chip or French fry) after harvest is crucial to allow for recovering from the increase of sugars and assure the suitability for processing by storing at appropriate temperatures. The future of sorting tubers based on internal chemical composition is growing in research and possible commercial systems might be available with the advancement on spectroscopic hardware (light sources, spectrophotometers) and pattern recognition methods (SIMA or soft independent modeling of class analogy, K-nearest neighbor or Knn, artificial neural network or ANN, support vector machines or SVM, decision trees) and finally with the appropriate arrangement of samples, light sources, detectors. Unlike other agricultural commodities (apple, pear, cucumber, etc.), the quite broad variation of shapes, sizes, and diverse uses of potato tubers, presents challenges for rapid and/or non-destructive technology application at points right after harvesting, handling, storage, or even after processing operations. The huge economic value associated with the potato industry obligates more research to develop cost effective, yet highly accurate, monitoring systems based on the current or future technologies to enhance food quality, safety, and human nutrition attributes.

CHAPTER 3 THE POTENTIAL USE OF VISIBLE/NEAR INFRARED SPECTROSCOPY AND HYPERSPECTRAL IMAGING TO PROCESSING-RELATED CONSTITUENTS OF POTATO TUBERS

(Rady, A.M., Guyer, D.E., Kirk, W., Donis-González, I.R. 2014. The potential use of visible/near infrared spectroscopy and hyperspectral imaging to predict processing-related constituents of potatoes. *Journal of Food Engineering*, Vol. 135: 11-25)

3.1 Introduction

In processing, tubers require a consistent internal composition that is maintained and achieved by monitoring important internal or external constituents that are strongly related, directly or indirectly, with product quality. For processing applications, dry matter which accounts for 18 to 26% of the tuber weight has an effect on frying process efficiency, product yield and oil absorption (Burton, 1989). Specific gravity is one of the most important physical properties of potato tubers and is strongly associated with dry matter content, which in turn is correlated with the yield of processed products, e.g. French Fries, chips, and dehydrated products (Kadam, 1991). Glucose is responsible for the undesirable browning color that follows the frying process and it dramatically affects the marketability of chips and other fried potato products. Such color is a result of the Maillard reaction which includes the interaction between an amino acid (asparagine) and the reducing sugars, glucose and fructose, (Mottram and Wedzicha, 2002). Moreover, acrylamide, discovered by the Swedish National Food Authority in 2002, (Zyzak et al., 2003; Stadler et al., 2002; Mottram and Wedzicha, 2002) is also formed during the frying process. The acrylamide single unit (monomer) is toxic to the nervous system, a carcinogen in laboratory animals and a possible carcinogen in humans. Consequently, monitoring glucose levels during storage is important to provide healthy, and high quality French fries, and chips. Sucrose level in potato tubers dedicated for processing is critical as it causes the unacceptable

sweetening flavor. The high level of sucrose is more likely to happen after the storage period, though this increase is cultivar-dependent. Soluble solids content is an important factor of the level of dissolved sugar in samples which indicates the ability of tubers to go to processing for chipping or French fry products. Although measuring this factor is relatively easy to perform using a refractometer, it is still an invasive method.

Primordial leaf count is an indication of the ability of tubers to grow and yield sprouts which is an important factor affecting the total crop yield. The leaf count is also an indication of the physiological status of a potato tuber which is important to monitor for seed potatoes (Kirk et al., 1985). The number of sprouts per seed tuber is determined by the size of tubers as well as the storage conditions (Allen et al., 1992). No significant work in the area of non-destructive evaluation of leaf count for potato tubers was found in the literature.

The importance of each of these constituents to food products, combined with the desire for highly correlated automated measurements, suggests the need for developing a rapid yet accurate, and possibly non-invasive, system that can be used as a trusted technique to monitor and help detect the postharvest properties of potato tubers.

Near infrared (NIR) spectroscopy has been known as a fast and non-destructive method to evaluate the internal and external quality factors for food products (Dufour, 2009; McClure, 2007; Shenk et al., 2001; Barton and Kays, 2001). Sukwon, et al. (2003), used NIR technology to develop a calibration model by which both percentage of dry matter and specific gravity of potato tubers can be calculated. The coefficient of determination of the specific gravity model was 0.87 with a correlation coefficient of 0.85; for dry matter percentage, the correlation coefficient was 0.82. Subedi, and Walsh (2009), demonstrated the advantage of using short-wavelength near-infrared spectroscopy (over the wavelength region 750–950 nm) to measure the

dry matter concentration of potato tubers; the correlation coefficient for the whole tubers was 0.85, the value increased to 0.95 for sliced tubers. NIR technology was also used for potatoes by Jeong et al. (2008), to estimate the sprouting capacity of tubers. Using the modified partial least square method (MPLS), the values of R ranged from 0.87 to 0.97 for the calibration models, and the values were 0.72 to 0.90 for the validation models. Hartmann, and Buning-Pfaue (1998), studied the use of NIR spectroscopy in measuring some constituents of peeled potato tubers. The diffuse reflectance mode was used in the wavelength range of 1100-2500 nm. Dry matter, starch, fructose, glucose, and sucrose were all measured using standard methods and the MPLS regression was used to build the models. The validation model had standard errors of 0.041%, 0.028%, and 0.037% with R^2 values of 0.70, 0.89, and 0.62 for glucose, fructose, and sucrose respectively.

Hyperspectral imaging systems (HIS) have been used in agriculture for two decades. HIS have several advantages, for example: (1) images the scene in hundreds of co-registered bands, (2) spectral resolution 10 X the order of multi spectral images (MSI), and (3) HIS have spectral bands that are contiguous and regularly spaced leading to continuous spectrum measured for each pixel (Kerekes and Schott, 2007). In addition, El Masry and Sun (2010), noted that HIS require minimal sample preparation; including non-destructive nature, and fast acquisition times with the capability of visualizing the spatial distribution of desirable constituents. HIS were studied in the area of defect detection and sorting operations as well as estimation of internal constituents in food materials (Molto et al., 2010; El Masry and Sun, 2010; Chao, 2010; Menesatti et al., 2010; Wang and El Marsy, 2010). Jun Qiao et al. (2005), studied the application of the hyperspectral imaging technique to estimate both the water content and the weight of potato tubers. The system was used to extract morphological features and spectral responses on

water content in tubers simultaneously. The wavelength range of 934-997 nm was found to be sensitive to the absorption band for predicting the water content in potato tubers. Results showed that the coefficient of correlation between the predicted and actual values of water content was 0.93 and 0.77 for training and validation, respectively. Lu and Peng (2006), used a hyperspectral imaging system to study hyperspectral scattering to estimate peach firmness; the Lorentzian distribution function was used to model the scattering profile, then multi-linear regression (MLR) along with cross-validation were used to build the calibration model which was then applied to a different validation set of data with coefficient of determination (R^2) of 0.67 to 0.77.

This research studies the objective of determining the potential of using VIS/NIR spectroscopy and hyperspectral imaging systems to estimate constituents in potato tubers that are important to the processing and seed industries.

3.2 Materials and Methods

3.2.1 Sample Collection, Handling, and Treatments

Two common cultivars of potatoes were used in the experiments; Frito Lay 1879 (FL) which is used in the chipping or crisping industry, and Russet Norkotah (RN), which is usually used as table-stock or ware for baking and boiling. The samples were obtained from commercial production fields in Southwest Michigan, USA. During September, 2008, there were two vine killing dates followed by two respective harvesting times for each cultivar, early and late, in an effort to obtain an extensive range of physiological characteristics. Tubers were cleaned and defective samples discarded, then all samples were stored at 7 °C for 3-4 weeks for initial curing and the first sampling was conducted at the end of this period. The samples were then stored in three temperatures; 7, 10, and 15 °C. Tubers were sampled after 20, 80, and 130 days of storage

to additionally aid in developing a strong and broad sample set. The experimental design and approach is depicted in Fig. 3.1

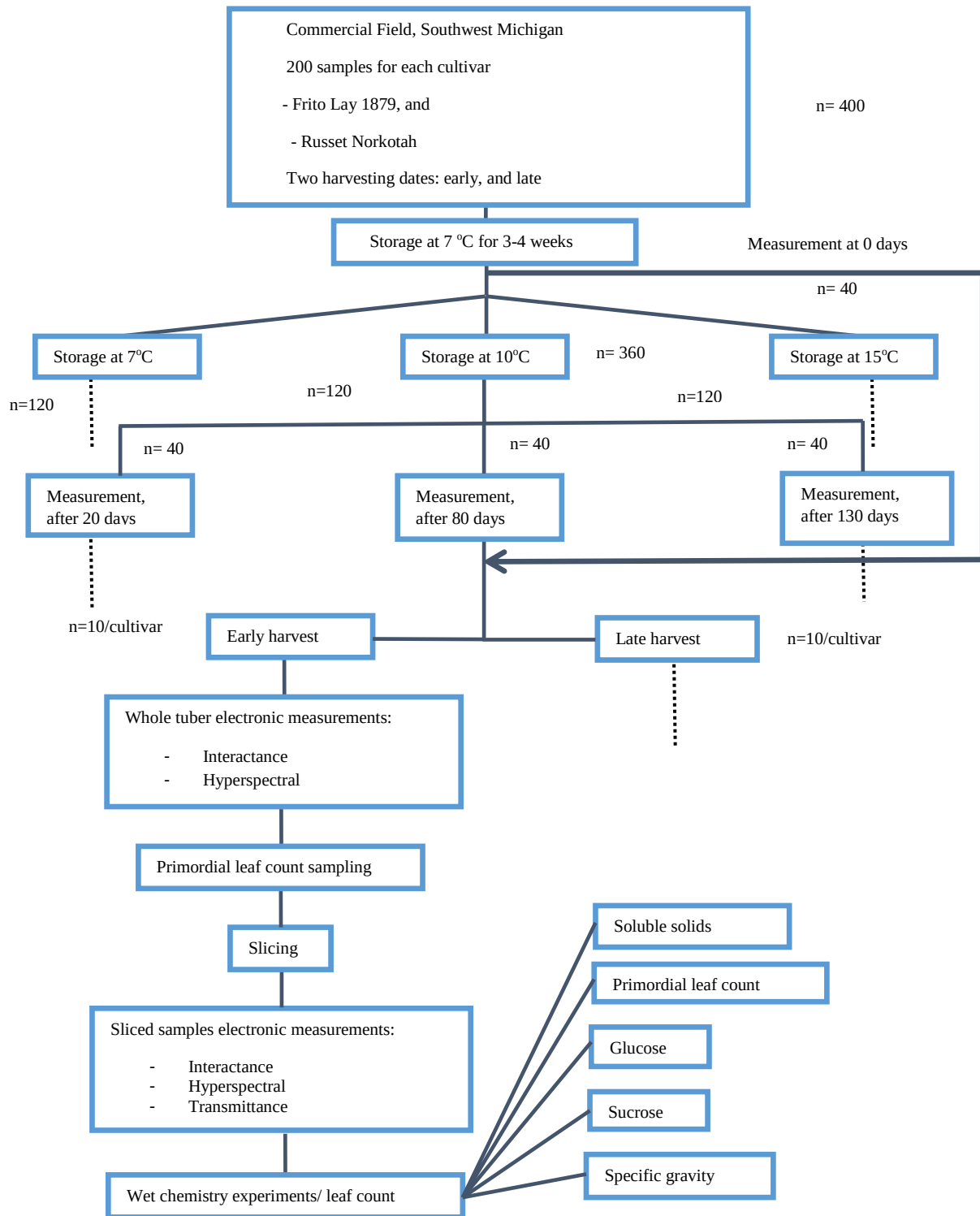


Figure 3.1. Flow chart of the experimental design to assess physiological status of potato tubers using visible/near infrared spectroscopy and hyperspectral imaging for Frito Lay 1879 and Russet Norkotah cultivars.

3.2.2 Electronic Measurement

3.2.2.1 Sample preparation

Either one or two types of samples were utilized for the rapid measurements of NIR transmittance, visible/NIR interactance, and visible/NIR hyperspectral reflectance. First, whole tubers, in which the sample was placed such that the light was directed to the middle area of the tuber, and the second type of sample comprised a 12.7 mm tuber slice obtained by cutting the tuber three times perpendicular to the longitudinal axis, starting from the stem end of the tuber. The measured slice was the third slice in the cutting routine and both sides of the slice were tested. Both the whole tuber and the sliced samples were used in the case of visible/NIR interactance and visible/NIR hyperspectral scattering modes, however, just the sliced samples were used in case of the NIR transmittance mode.

3.2.2.2 VIS/NIR interactance mode

In the interactance mode, light photons illuminated the sample by a probe with a concentric outer ring of illumination and an inner receptor (Fig. 3.2). In this case, the overall probe was in contact with the sample surface, and a foam-sealing ring separated the ring of light and the detector, so only the light interacting within the sample was measured. The system used for interaction experiments contained an Ocean Optics fiber optic spectrometer (model No. USB 4000, Ocean Optics, Inc., Dunedin, FL, USA) with an optical resolution of 0.3 nm (FWHM), and with a 200 μm diameter fiber optic, Oriel radiometric power supply with a maximum power of 250 watt (model No.68931, Oriel Inst., Irvine, CA, USA), and Oriel light source (model No. 66881, Oriel Inst., Irvine, CA, USA) with the same maximum power and the wavelength measurement range between 446 to 1125 nm, covering both visible and NIR fields. With this configuration, the incident light represented a circle with a diameter of 24.7 mm. The

interactance experiment was conducted on both sliced and whole samples. The interactance for each sample was normalized using Teflon[®] as a reference material, and the relative interactance was calculated using equation 3.1 as follows:

$$\text{Relative Interactance} = \frac{\text{intensity of sample interactance} - \text{intensity of background interactance}}{\text{intensity of reference interactance} - \text{intensity of background interactance}} \quad (3.1)$$

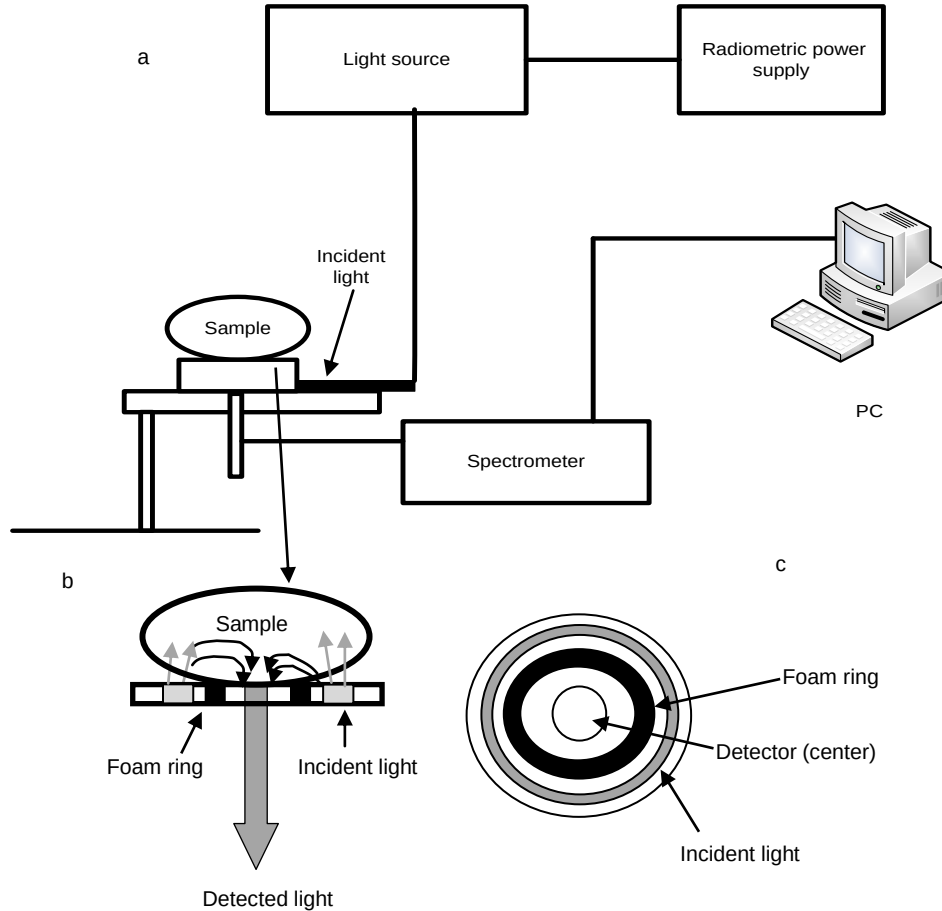


Figure 3.2, a. Schematic representation of VIS/NIR interactance mode used to predict constituents for two potato cultivars, b. Light path representation, c. End view of probe.

3.2.2.3 VIS/NIR hyperspectral mode

The hyperspectral system in this project was used to capture diffuse scattered light from both the whole and sliced samples in the range of 400 to 1000 nm, thus covering visible and near NIR bands. The system used to study the samples under hyperspectral reflectance mode contained a Hamamatsu dual mode cooled CCD camera (model No.C4880, Hamamatsu Photonics, Hamamatsu, Japan) along with an Oriel power supply (model No.69931, Oriel Inst., Irvine, CA, USA), an Oriel digital exposure controller (timer) (model No.68945, Oriel Inst., Irvine, CA, USA), Agilent DC power supply (model No.65423A, Agilent Tech., Santa Clara, CA, USA), and Oriel light source (model No. 66881, Oriel Inst., Irvine, CA, USA) that contained a quartz tungsten halogen lamp. (Fig. 3.3a). The imaging spectrograph acquired spectral information by working in the point scan mode where the collimated light was dispersed from the sample into different wavelengths by a prism-grating-prism configuration while keeping spatial information at the same time. The CCD camera detected the dispersed light signals and created a 2-D image, 256 X 256 pixels, with the horizontal axis representing the spatial values and the wavelength values were recorded on the vertical axis. The sample holder was moved with a motor controlled stage and allowed consistent height between sample and detector/light source and for multiple scanning points for each sample. The distance between two successive scans was adjusted at 1 mm, and a total number of 10 images (scans) were acquired for each sample. Thus a set of images was a data-cube, representing spectral information of a 9 mm longitudinal distance along the sample. The scattering behavior of light in a sample was shown in Fig. 3.3b. Light radiation penetrated the sample surface and scattered outward through the tissue, and the diffuse reflected light was captured by the CCD camera and spectrograph as a line scan. The scanning line was 1.5 mm apart from the incidence center. Fig. 3.4, a. shows a

sample of the 2-D image resulting from one scan of the sample, and fig. 3.4 b, and c show examples of spectral profiles from different spatial locations and at different wavelengths.

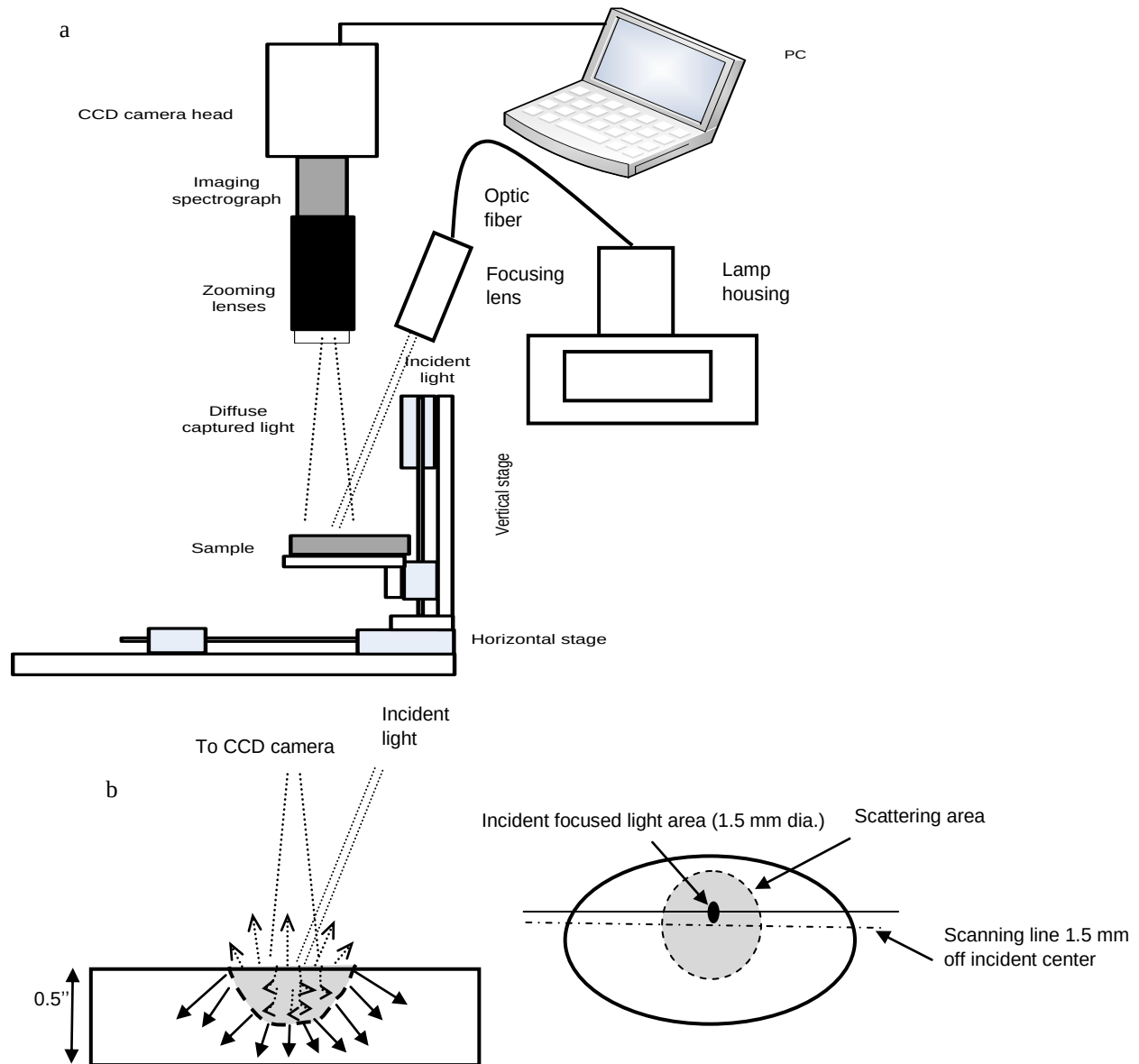


Figure 3.3, a. Schematic representation of VIS/NIR hyperspectral reflectance mode used to predict constituents for two potato cultivars, b. Light scattering in sample and scanning configuration.

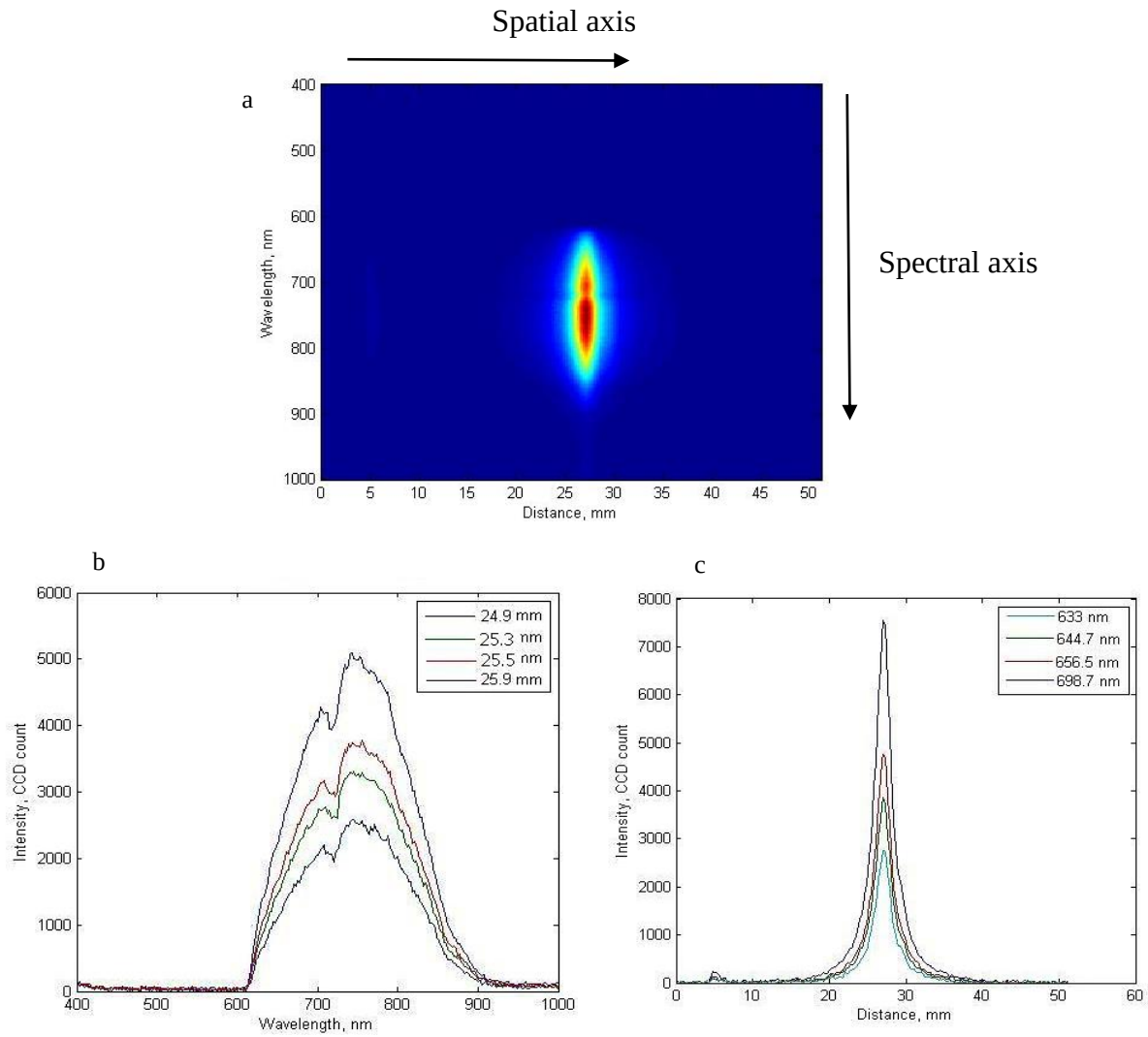


Figure 3.4, a. Hyperspectral scattering image, with different colors representing light intensity of a potato slice, b. Spectral profiles from different spatial locations represented by different colors, c. Spectral profiles from different wavelengths represented by different colors.

3.2.2.4 NIR transmittance mode

In the transmittance mode, the incident light vertically penetrates the sample surface and a portion of the incident light passes through the sample tissue to the other side with information about the internal composition of tubers (Chen, 1978). Both the light source probe tip and the detector tip were approximately 3 cm from the sample lower and upper surfaces respectively.

An InGaAs spectrometer (model No. NIR512L-1.7T1, Control Development, Inc., South Bend, IN, USA) with spectral resolution of 3.25 nm FWHM and linear dispersion of 1.625 nm/pixel was used in the transmittance mode along with an Oriel radiometric power supply with a 300 watt maximum power (model No.68931, Oriel Inst., Irvine, CA, USA), and an Oriel light source (model No. 66881, Oriel Inst., Irvine, CA, USA) that has 250 watt maximum power, and with a quartz tungsten halogen lamp. Only the sliced samples, with 0.5'' (12.7 mm) thickness each, were used in the transmittance experiments with the sample area covered by the detector having a diameter of 1'' (25.4 mm). The calculation of the relative transmittance was done over the NIR wavelength range between 900-1685 nm in the same way as in the calculation of relative interactance. A schematic diagram of the transmittance system used in the experiment was represented in Fig. 3.5.

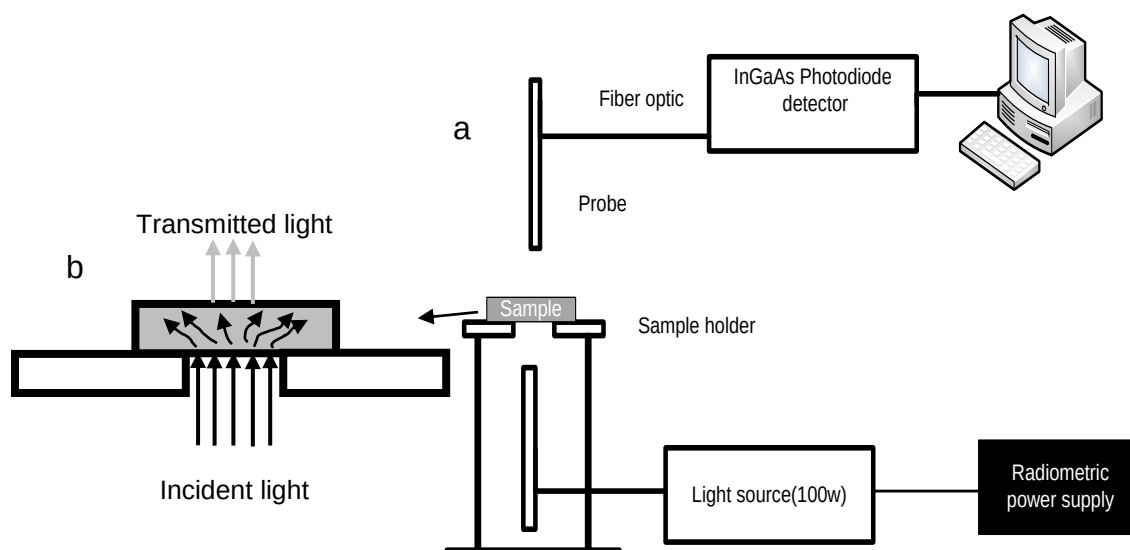


Figure 3.5, a. Schematic representation of NIR transmittance mode configuration and system components, b. Light path representation with scattering in the sample and the detected transmitting light.

3.2.3 Constituent (Reference) Measurement

3.2.3.1 Measurement of glucose and sucrose

3.2.3.1.1 Extraction of juice

The objective of this destructive process was to provide a validation/reference for the amount of sucrose and glucose in the tuber or piece of tuber that was subjected to the electronic measurements. The standard method used to estimate glucose and sucrose is the enzymatic method, using glucose oxidase and paraoxidase enzymes. Directly after conducting the electronic measurements for each whole tuber, and then for the sliced sample, the sample was put in a labeled plastic bag, and stored in a foam box containing ice to maintain the tubers in a fresh state and minimize any chemical changes during the performing of the electronic measurements for the additional samples. Each tuber was then put in a Juicerator 6001, 500 watt (ACME Supreme, New Hartford, CT., USA) to extract the juice from the tuber by centrifugal force at 3600 RPM

and using a paper filter that was placed around the inner surface of the Juicerator. The juice was transferred with a pipette to a polystyrene tube with cap and then stored at -20 °C to reduce any variation of constituents and allow subsequent use and analysis of the juice at a later time.

3.2.3.1.2 Chemical estimation of glucose and sucrose

Using the Megazyme sucrose/D-glucose assay procedure (Megazyme International Ireland Ltd, Wicklow, Ireland), the ratio of each of glucose and sucrose, gram per 100 gram fresh tuber weight, was measured. Tubes containing frozen juice sample were thawed at 18 °C. As the concentration of both glucose and sucrose for RN was higher than for FL, 100 µL of juice was transferred to each of four glass test tubes from the FL samples, whereas for RN, the 100 µL volume consisted of 10 µL juice diluted by 90 µL of distilled water. To estimate the glucose ratio, 100 µL of sodium acetate buffer, 2M, was added to two tubes, and to estimate sucrose ratio 100 µL of β-fructosidase (invertase) diluted by sodium acetate buffer was added to the other two tubes. The tubes were incubated in a water bath set at 50 °C for 20 minutes, then 1500 µL of glucose determination reagent (GOPOD reagent) was added and the samples were incubated under the same conditions in the water bath for 20 minutes. The content of each tube was transferred to a 96 well (200 µL) ELISA plate and the absorbance of the solution was measured at 510 nm in a spectrophotometer against both the blank sample of 100 µL distilled water which was prepared using the same procedure and the control sample of 50 µL of D-glucose standard + 50 µL distilled water. The D-glucose, or dextrose, and sucrose concentrations were then calculated using equation 3.2 and 3.3 respectively:

$$\text{D-glucose (g/g fresh weight)} = \Delta A \times F \times 0.005 \times 1/10 \quad (3.2)$$

$$\text{Sucrose (g/g fresh weight)} = (\Delta B - \Delta A) \times F \times \text{Dilution} \times 0.0095 \times 1/10 \quad (3.3)$$

Where:

ΔA : represents the GOPOD absorbance for D-glucose

ΔB : represents the GOPOD absorbance for sucrose

F: is a factor used to convert from absorbance to μg for 100 μg of D-glucose (100/absorbance for 100 μg D-glucose); and

Dilution: 1 in case of Frito Lay 1879 and 10 in case of Russet Norkotah

0.1: Unit conversion factor to convert from g/L into g/100g or % fresh weight

3.2.3.2 Measurement of soluble solids

The soluble solid content is the concentration of the solid particles in a solution and it usually refers to the sugar concentration but without expressing the sugar type, Thus, one can't depend only on the soluble solids as an indication of sugar concentration though the advantage of rapid assessment of such constituent exists using modern digital refractometers with the Brix unit. Soluble solids concentration was measured using a Palette digital refractometer (model No. PR-101, ATAGO Co. LTD, Bellevue, Washington, USA) by dripping juice on the device prism and reading the displayed Brix units.

3.2.3.3 Measurement of specific gravity

The specific gravity was indirectly measured using the relationship with the dry matter mentioned by Kellock (1995). Such relationship is as follows:

$$SG = 0.0053 * DM + 0.960574 \quad (3.4)$$

Where:

SG: is the specific gravity, g/cm^3 ; and

DM: is the percentage dry matter (dry matter weight divided by the total tuber weight X 100)

After juicing, the filter-collected solids from the samples were placed inside a drying oven at 100 °C for 24 hours and weighed to calculate dry matter (DM) and the SG was calculated using equation 3.4.

3.2.3.4 Measurement of primordial leaf count

The number of leaf primordia within the developing sprouts gives an indirect measurement of tuber maturity or physiological age (Kirk et al., 1985). Counts of leaf primordia were conducted by taking samples of eyes from each tuber (n=3) before juicing the tuber for future estimation of glucose and sucrose. The samples were chosen from the apical end of the tuber. Briefly, the sprouts were stored in 5 ml Eppendorf tubes in an ethanol:acetone solution (1:1) until used. Sprouts were mounted on slides and examined at 10x magnification under a dissecting Olympus microscope (model No. ZT40, Olympus Corp., Tokyo, Japan). Leaf initials were removed sequentially from the outside to inside of the sprout using a scalpel until the apical dome was exposed. Leaf primordium counts were obtained for the three eyes and then the average was taken and considered the primordial leaf count per tuber.

3.2.4 Partial Least Squares Regression (PLSR)

Partial least squares regression (PLSR), also called projection to latent structures by means of partial least squares, is a powerful linear regression method that is insensitive to collinear variables and tolerant to large numbers of variables (Varmuza and Filzmoser, 2009).

3.2.4.1 Pretreatment of the spectra data

When the signals are acquired from a set of samples it may be necessary to pretreat data before building the calibration model (Christy and Kvalhiem, 2007). This is because the original data sometimes contains unwanted spectral variation and baseline shifts that may be a result of light scattering from samples, the poor reproducibility of NIR spectra due to path length variation, variation of the sample conditions (temperature, particles' sizes), and various noise resulting from detector, A/D convertor, and other electric components in the system. Preprocessing methods depend either on abstractly mathematical concepts, or previous knowledge of the chemical–physical background of the data and the discussed problem (Varmuza and Filzmoser, 2009).

The sequence of processing was in two stages for the spectra data. The first stage was a primary processing method that may be in addition to the option of non-preprocessing. This stage included absolute value, autoscaling, baseline, weighted baseline, smoothing with first derivative, smoothing with second derivative, normalization, generalized least squares weighting, standard normal deviate (SNV) correction, multiplicative signal correction (MSC), group scale, and median center. The second stage of preprocessing, that treats the first stage treated data, is the one included in the PLSR algorithm that is conducted by Eigenvector (Eigenvector Research, Inc. WA, USA) using the platform of Matlab[®] software (version 7.5.0.342, MathWorks, Natick, MA, USA) and that is either the mean center method, multiplicative scattering correction (msc), or orthogonal signal correction (osc) (Wise et al., 2006). A flow chart of the preprocessing steps conducted for the spectra data was shown in Fig. 3.6a.

3.2.4.2 Pretreatment of the reference data

Transformation of the reference data (the dependent variable in the regression model) was conducted with the aim to get the constituents' distribution as uniform as possible. Such transformation includes the log and power transformation, with 2.0 as the exponent, in addition to using the non-transformed data to study the effect of constituents' values transformation. The preprocessing steps for reference data were clarified in Fig. 3.6b.

Calibration and validation sets of data were formed such that the calibration set contained 75% of the data and the validation set contained 25% of the data. The cross validation technique (leave-one-out) was used to get the best calibration model based on the minimum mean square of error for calibration for cross validation (RMSECcv) and the calibration model was subsequently applied to the validation or prediction set.

The results presented later are the best from the different preprocessing methods based on the correlation coefficient (R), root mean square error of prediction for validation set (RMSEP), and the RPD value (the standard deviation of the reference data divided by the RMSEP). In general, root mean square error, either for calibration or validation, is calculated using the following equation:

$$RMSE = \left[\frac{\sum_{i=1}^N (Y_i - \hat{Y}_i)^2}{N} \right]^{1/2} \quad (3.5)$$

Where:

N: number of samples

Y_i : actual value of reference (constituent) for sample i ; and

$\hat{Y}_i$: predicted value of reference (constituent) for sample i

Also, coefficient of correlation (R) is calculated using equation 3.6 as follows:

$$R = \frac{\sum_{i=1}^N (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^N (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^N (Y_i - \bar{Y})^2}} \quad (3.6)$$

Where:

X_i : Relative intensity value for sample i

$\bar{X}$: Average of relative intensities for data set; and

$\bar{Y}$: Average of reference values for data set

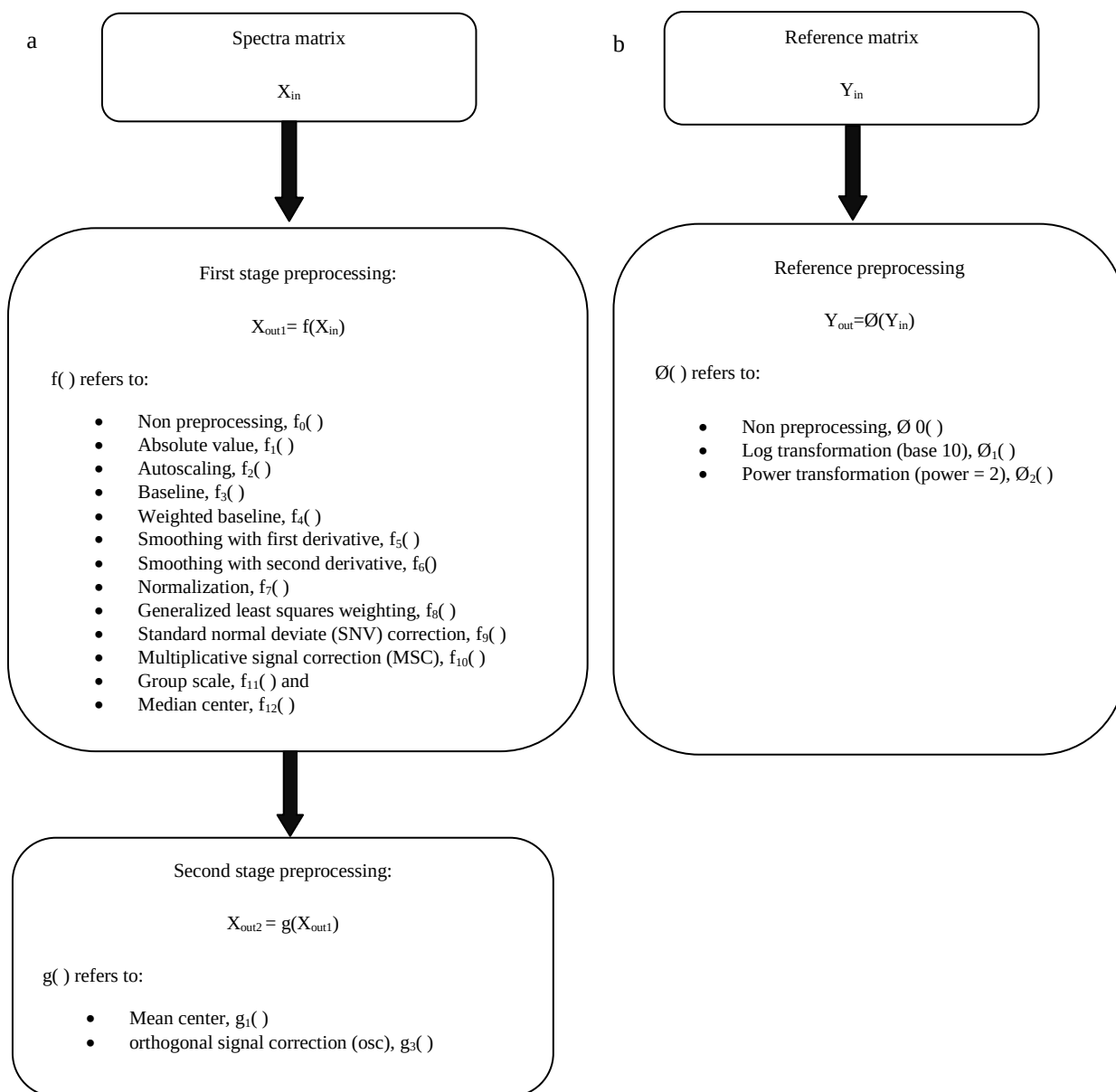


Figure 3.6. Flow chart of preprocessing methods used to pretreat spectra, a, and reference, b, data before building calibration and then prediction models using PLSR with cross validation to predict constituents for two potato cultivars.

3.3 Results

3.3.1 Constituents' Distributions

The results of reference analysis, after discarding the logical outliers for each constituent that are located outside the expected range, based on literature (Storey, and Davis, 1992), are shown in Fig. 3.7a to 3.7e (n=200/cultivar). Some figures also show statistical outliers such as: glucose, sucrose, and primordial leaf count. Following the fact that RN has higher sugars levels (glucose, and sucrose) than FL 1879, it's clear from Fig. 3.7. a, and b the difference between the two cultivars in these sugars. However, for specific gravity, primordial leaf count, and soluble solids there was no significant difference between the two cultivars.

3.3.2 Spectra for Different Modes

3.3.2.1 Interactance mode

The mean signals acquired from interactance mode, for sliced samples in the case of glucose and sucrose in two ranges for both cultivars, were shown in Fig. 3.8a-d. The thresholds were chosen as the median value. For FL, the thresholds for glucose, and sucrose were 0.02%, and 0.05 % respectively, whereas those values for RN were 0.2%, and 0.07% respectively. For FL, there was no clear difference between the mean spectra in both glucose and sucrose. However, the difference was more evident in the case of RN for both sugars which is a result of the higher levels of sugars in the case of RN compared with FL. The same trend of mean spectra was found for whole tubers (Fig. 3.9) although there was a slight difference for FL in the case of glucose compared with sliced samples.

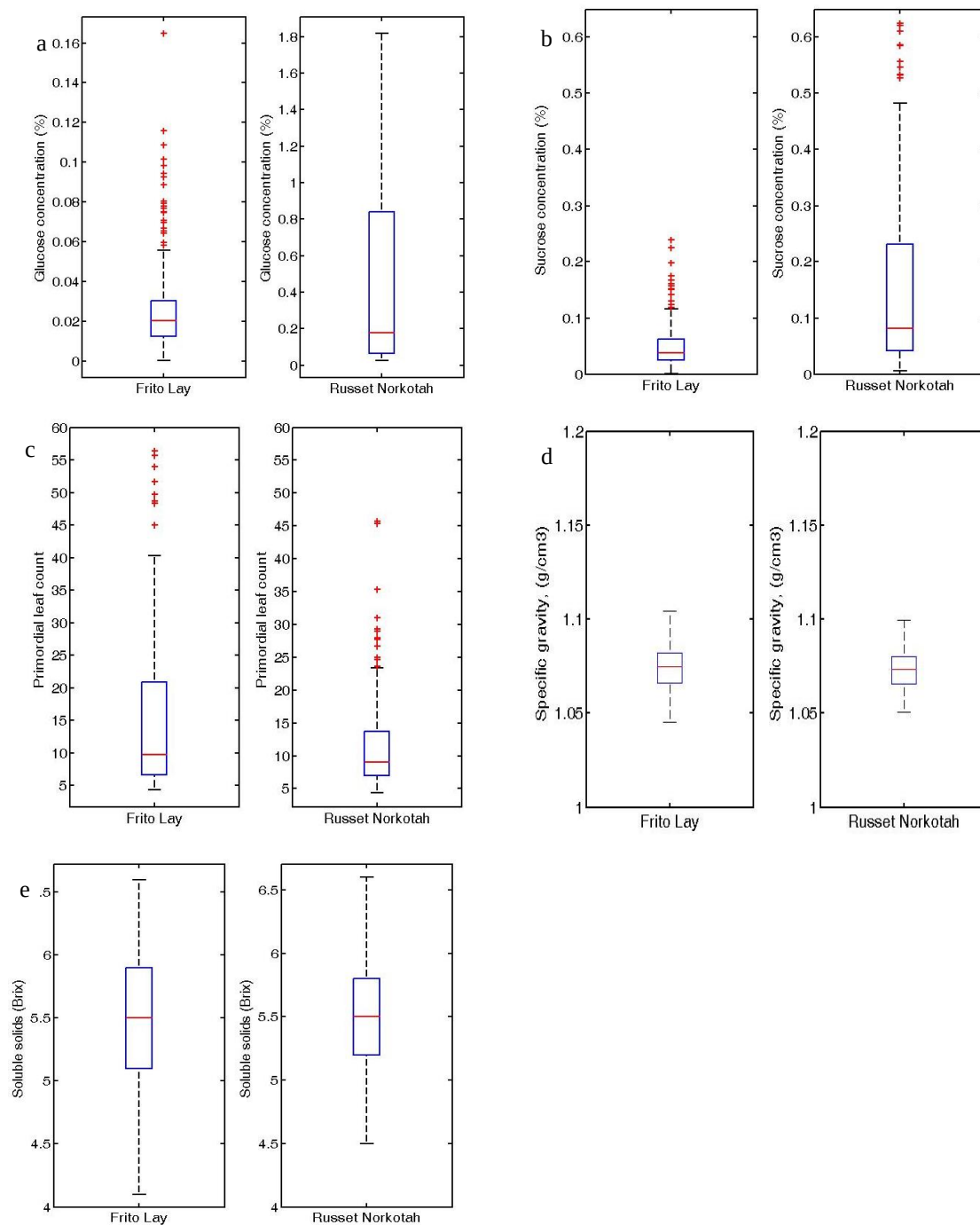


Figure 3.7. Data distributions of the physiological variables measured; a. Glucose concentration % (note change in range of the values for the two cultivar types), b. Sucrose concentration (%), Primordial leaf count (number of leaves per sprout), d. Specific gravity (g/cm³), e. Soluble solids (Brix scale).

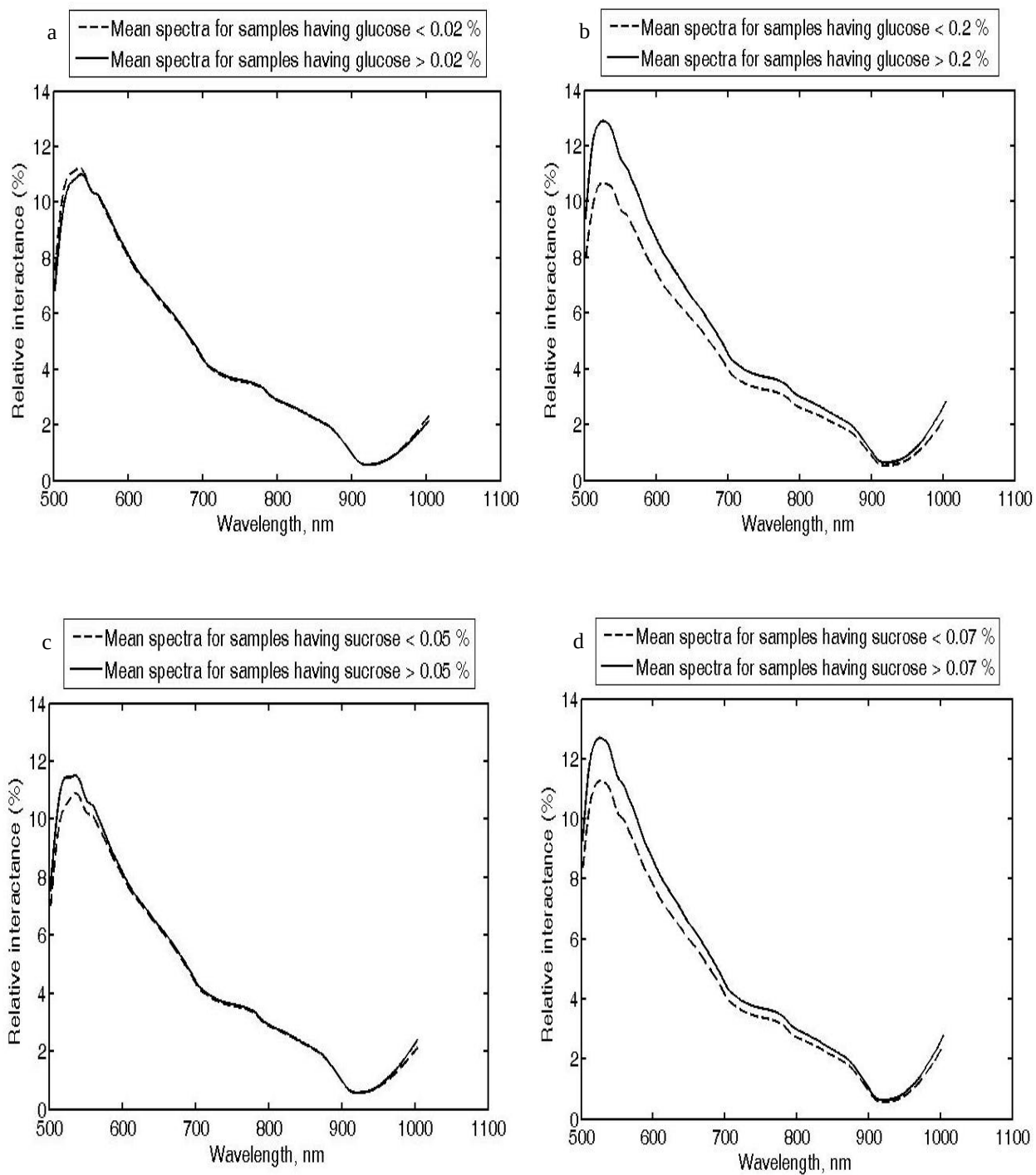


Figure 3.8. Mean relative intertance for two sugar groupings for sliced samples, a. Frito Lay 1879: glucose, b. Russet Norkotah: glucose, c. Frito Lay 1879: sucrose, and d. Russet Norkotah: sucrose.

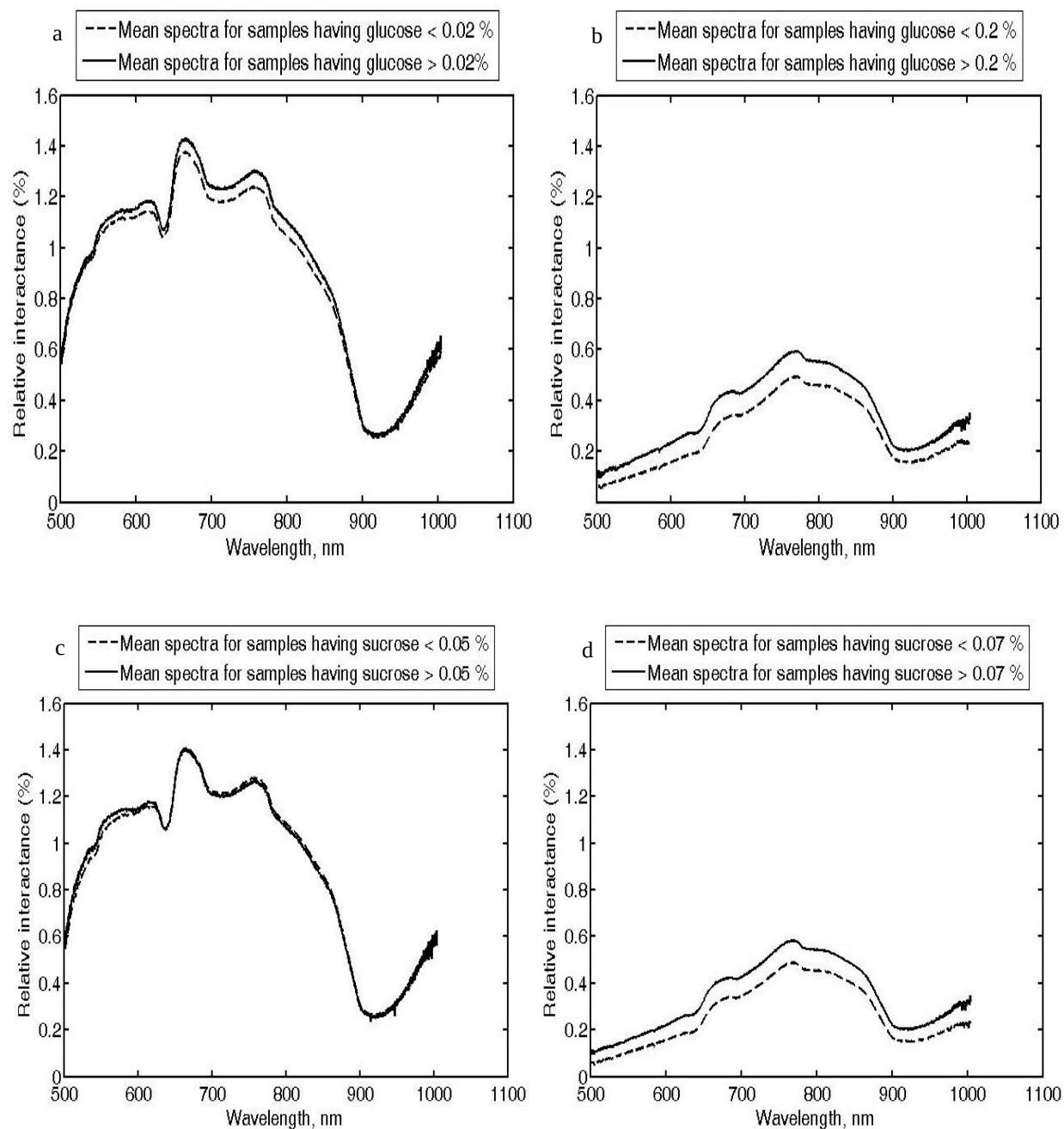


Figure 3.9. Mean relative interactance for two sugar groupings for whole tubers, a. Frito Lay 1879: glucose, b. Russet Norkotah: glucose, c. Frito Lay 1879: sucrose, and d. Russet Norkotah: sucrose.

3.3.2.2 Hyperspectral imaging mode

The mean reflectance spectra for the hyperspectral data were collected for sliced and whole samples for both cultivars. To extract the mean reflectance for each image, all wavelengths in the range 400 to 1000 nm were used. The spectra were normalized by the Teflon[®] reference average reflectance spectra. The mean reflectance spectra for two ranges for glucose for both cultivars in the case of sliced samples were shown in Fig. 3.10a, and b. Both cultivars had an absorption band at 837 nm which is likely related to the hydrocarbon group C-H, aliphatic with another one at 880 which is possibly due to aromatic associated C-H group (Workman and Weyer, 2008). Moreover, difference between the two sugar classes is higher in RN than FL for glucose. For sucrose, the same absorption band was yielded while the difference between two classes in the case of FL is higher than RN. In the case of whole tubers, no significance difference between the two sugar groups was found except in the case of sucrose for FL (Fig. 3.11). In general, the mean relative reflectance overall is less for the whole tubers than for the sliced samples for both cultivars with the note that the skin effect is more obvious for RN than FL in the case of whole tubers due to the thicker skin for RN compared with FL.

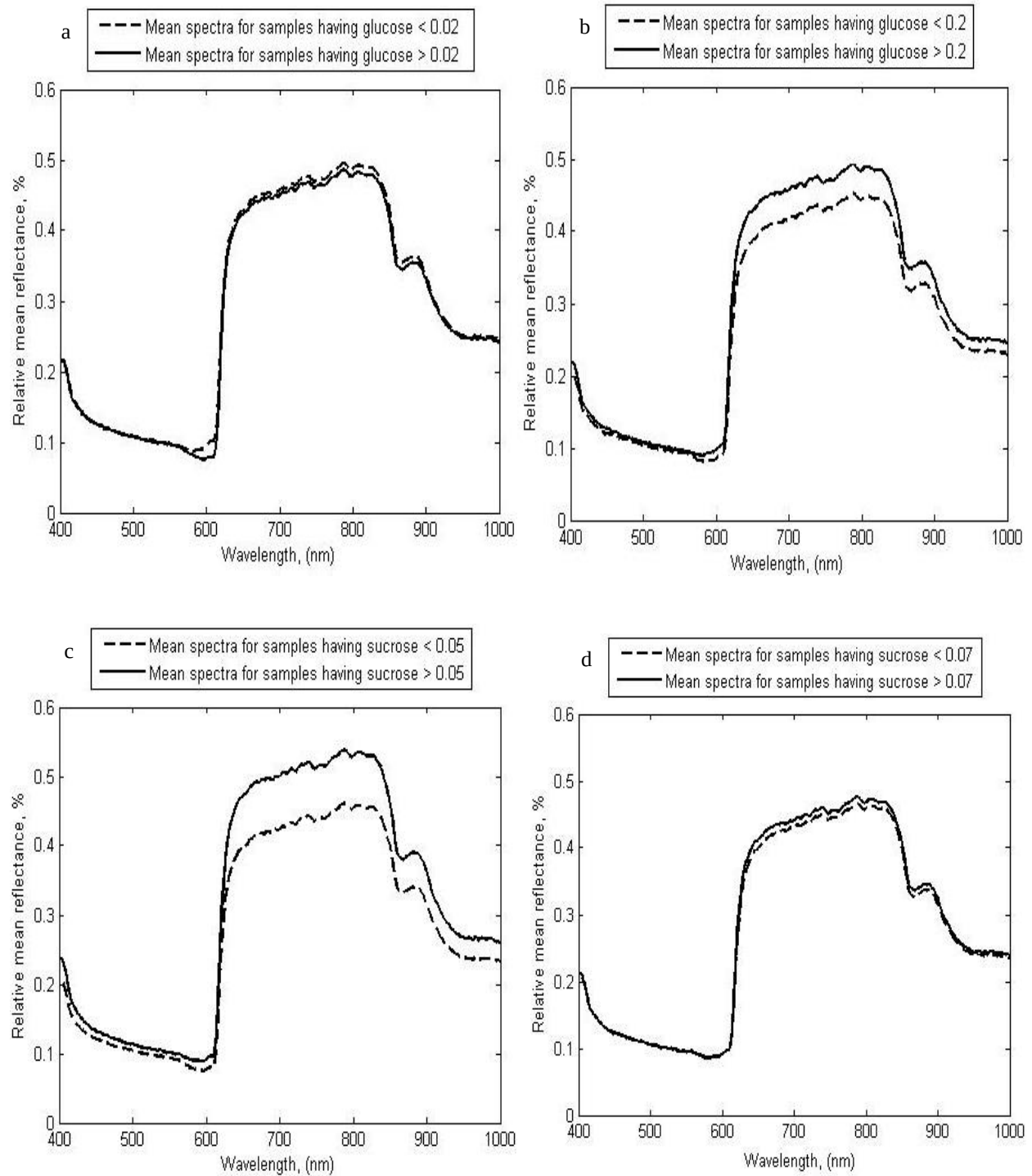


Figure 3.10. Mean relative reflectance for two sugar groupings for sliced samples, a. Frito Lay 1879: glucose, b. Russet Norkotah: glucose, c. Frito Lay 1879: sucrose, and d. Russet Norkotah: sucrose.

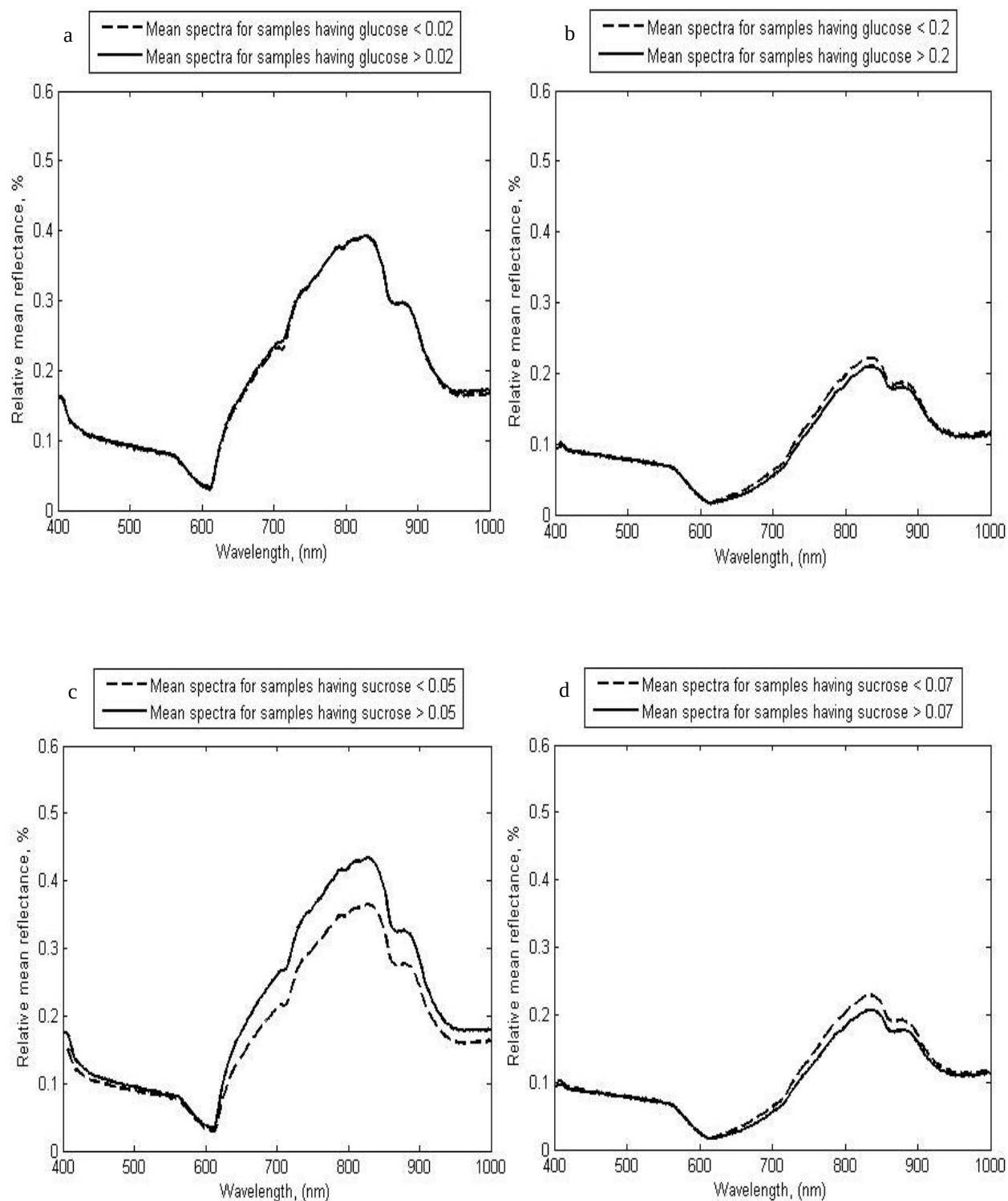


Figure 3.11. Mean relative reflectance for two groupings for whole tubers, a. Frito Lay 1879: glucose, b. Russet Norkotah: glucose, c. Frito Lay 1879: sucrose, and d. Russet Norkotah: sucrose.

3.3.2.3 Transmittance mode

The mean relative transmittance signals acquired from both cultivars for sliced samples is shown in Fig. 3.12 for the wavelength range of 900 to 1685 nm, for two ranges of glucose and sucrose as explained in section 3.3.2.1. There are peaks at 1200 nm and 1430 nm, in all cases that are suspected as systematic error from instrumentation because of their consistency and repeatability. Slight differences were observed between the mean spectra of the different ranges for the glucose and sucrose for FL, whereas, the difference is more visible for RN, and again the possible reason for this is the higher levels of sugars for RN.

The spectral plots for each electronic mode helped to interpret the performance of prediction models yielded from PLSR based on the idea that if the optical mode is capable of acquiring different values of chemical constituents in differentiating between samples (the difference between the two classes of curves), there will be more likelihood to obtain high prediction models. Other constituents' plots (not shown) resulted in similar findings.

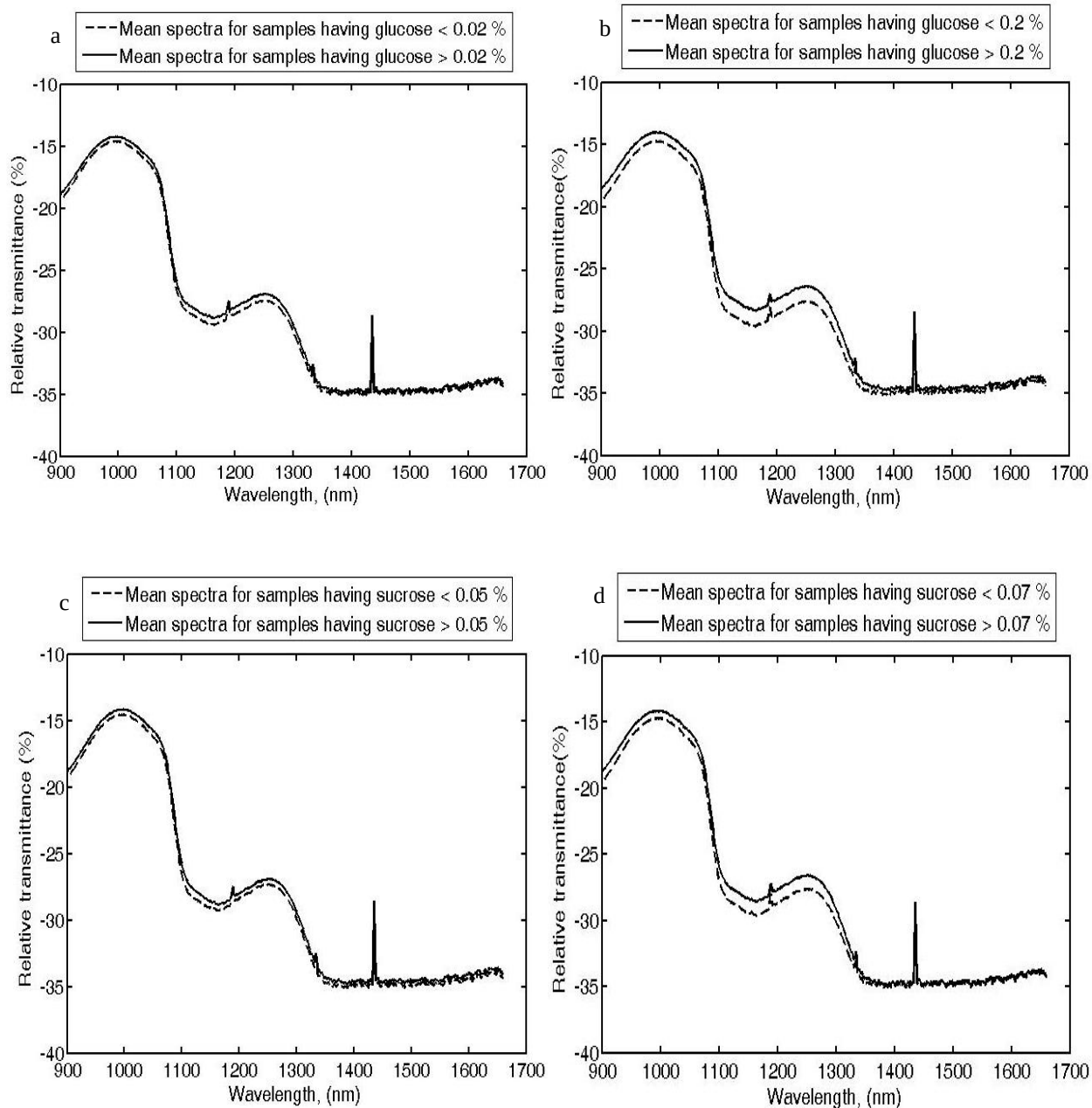


Figure 3.12. Relative transmittance for two sugar groupings for sliced samples, a. Frito Lay 1879: glucose, b. Russet Norkotah: glucose, c. Frito Lay 1879: sucrose, and d. Russet Norkotah: sucrose.

3.3.3 Partial Least Squares (PLSR) Results

3.3.3.1 Results for interactance mode

The responses for the interactance mode for sliced samples for each potato constituent are shown, with the best preprocessing sequence for spectra and for the reference data, in Table 3.1. The leaf primordial count prediction model for FL yielded R and RPD values of 0.95 and 3.29 respectively. The same model values for RN were 0.90 and 2.19 respectively. The glucose prediction model also had strong correlation for RN with R and RPD values of 0.95 and 3.12 and FL glucose values of 0.90 and 2.14 respectively. The sucrose prediction models were somewhat weaker than glucose for FL with correlations of R and RPD of 0.81 and 1.63 in contrast to RN for such which were much lower at 0.50 and 1.13 respectively. The other two constituents, specific gravity and soluble solids, did not yield as encouraging correlations as did the other three constituents.

In most constituents, correlation for whole tubers was less than that for sliced samples for interactance mode. For glucose, R and RPD values for FL of 0.88 and 1.78 respectively and 0.79, and 1.60 for RN (Table 3.2). Correlation for leaf count was found to be less than that for sliced samples for FL with values for R and RPD of 0.89 and 2.22 and 0.77 and 1.50 for RN respectively. Sucrose prediction for FL was somewhat stronger than for sliced samples with correlation metrics R and RPD values of 0.81 and 1.64 in contrast to RN that yielded weaker performance than sliced samples. Specific gravity prediction models for both cultivars showed less correlation than leaf count and glucose with best results obtained for sliced samples with R and RPD values of 0.37 and 1.06 for FL and 0.51 and 1.08 for RN. Other constituents showed poorer correlation which was the same trend as with sliced samples.

Table 3.1. PLSR results for predicting some potato constituents using VIS/NIR interactance (sliced samples) for Frito Lay 1879 and Russet Norkotah cultivars.

Cultivar _{Constituent}	Preprocessing ^a	Calibration			Prediction		
		R _{cal}	RMSEC _{CV}	LVs	R _{pred}	RMSEP	RPD
FL _{GL}	A ₆ , B ₁ ; C ₂	0.93	0.0553	10	0.90	0.0515	2.14
FL _{LC}	A ₀ , B ₁ ; C ₁	0.96	0.1979	13	0.95	0.2212	3.29
FL _{SG}	A ₁₂ , B ₁ ; C ₀	0.68	0.0099	12	0.61	0.0119	1.27
FL _{SS}	A ₆ , B ₁ ; C ₀	0.67	0.4378	10	0.55	0.4006	1.18
FL _{SU}	A ₅ , B ₁ ; C ₀	0.86	0.0490	10	0.81	0.0439	1.63
RN _{GL}	A ₇ , B ₁ ; C ₂	0.96	0.0858	15	0.95	0.0786	3.12
RN _{LC}	A ₇ , B ₁ ; C ₂	0.94	0.1625	13	0.90	0.1632	2.19
RN _{SG}	A ₀ , B ₃ ; C ₀	0.73	0.0090	10	0.54	0.0083	1.15
RN _{SS}	A ₁₀ , B ₃ ; C ₀	0.37	0.3970	4	0.37	0.3191	1.08
RN _{SU}	A ₇ , B ₃ ; C ₁	0.79	0.9792	2	0.50	1.0273	1.13
^a A_x: First stage spectra preprocessing. A ₀ : No preprocessing. A ₁ : Absolute value. A ₂ : Autoscaling. A ₃ : Baseline. A ₄ : Weighted baseline. A ₅ : 1 st derivative. A ₆ : 2 nd derivative. A ₇ : Normalization. A ₈ : Generalized least square weighting. A ₉ : Standard normal variate (SNV). A ₁₀ : Multiplicative signal correction (MSC). A ₁₁ : Group scale. A ₁₂ : Median center. B_x: Second stage spectra preprocessing. B ₁ : Mean center. B ₂ : Orthogonal signal correction. C_x: Reference data preprocessing. C ₀ : No reference transformation. C ₁ : Log reference transformation. C ₂ : Power reference transformation.							

Table 3.2. PLSR results for predicting some potato constituents using VIS/NIR interactance (whole tubers) for Frito Lay 1879 and Russet Norkotah cultivars.

Cultivar _{Constituent}	Preprocessing ^a	Calibration			Prediction		
		R _{cal}	RMSEC _{CV}	LVs	R _{pred}	RMSEP	RPD
FL _{GL}	A ₄ , B ₁ ; C ₂	0.96	0.0636	12	0.88	0.0620	1.78
FL _{LC}	A ₇ , B ₁ ; C ₁	0.99	0.3055	18	0.89	0.3285	2.22
FL _{SG}	A ₀ , B ₁ ; C ₀	0.45	0.0109	6	0.37	0.0143	1.06
FL _{SS}	A ₁₀ , B ₁ ; C ₀	0.19	0.4812	1	0.04	0.4834	0.98
FL _{SU}	A ₁₂ , B ₁ ; C ₀	0.89	0.0501	6	0.81	0.0436	1.64
RN _{GL}	A ₉ , B ₁ ; C ₂	0.88	0.1410	10	0.79	0.1529	1.60
RN _{LC}	A ₄ , B ₁ ; C ₀	0.91	0.4183	18	0.77	0.3560	1.50
RN _{SG}	A ₁₂ , B ₁ ; C ₀	0.72	0.0105	11	0.51	0.0089	1.08
RN _{SS}	A ₀ , B ₃ ; C ₀	0.46	0.4146	6	0.25	0.3431	1.01
RN _{SU}	A ₄ , B ₁ ; C ₀	0.71	0.1642	11	0.26	0.2051	0.97

^a See table 3.1 footnote.

3.3.3.2 Results for hyperspectral reflectance mode

The results of PLSR for hyperspectral reflectance for sliced samples showed strong correlation only for FL in the case of leaf count with R and RPD values of 0.94 and 2.92 respectively (Table 3.3). However, RN showed less correlation than interactance mode with R and RPD values of 0.70 and 1.41. Both cultivars showed less correlation for glucose prediction model with R and RPD values of 0.64 and 1.25 respectively for FL, and 0.74 and 1.49 for RN.

PLSR model results for hyperspectral reflectance for the whole tubers demonstrated significant lower correlation than the sliced samples for leaf count for both cultivars with R and RPD values of 0.47 and 1.14 respectively for FL and 0.43 and 1.10 respectively for RN (Table 3.4). The glucose model for RN also demonstrated low correlation with R, and RPD values of: 0.38 and 0.93 respectively, and 0.52 and 1.19 for FL. Sucrose, specific gravity and soluble solid content prediction models also showed weak correlations.

Table 3.3. PLSR results for predicting some potato constituents using VIS/NIR hyperspectral imaging (sliced samples) for Frito Lay 1879 and Russet Norkotah cultivars.

Cultivar _{Constituent}	Preprocessing ^a	Calibration			Prediction		
		R _{cal}	RMSEC _{CV}	LVs	R _{pred}	RMSEP	RPD
FL_{GL}	A ₁₂ , B ₁ ; C ₂	0.87	0.1024	6	0.64	0.0880	1.25
FL_{LC}	A ₉ , B ₁ ; C ₁	0.96	0.3256	4	0.94	0.2492	2.92
FL_{SG}	A ₉ , B ₁ ; C ₀	0.27	0.0112	2	0.26	0.0146	1.04
FL_{SS}	A ₅ , B ₁ ; C ₀	0.36	0.4702	4	0.14	0.4804	0.99
FL_{SU}	A ₅ , B ₁ ; C ₀	0.78	0.0636	12	0.62	0.0580	1.23
RN_{GL}	A ₁ , B ₁ ; C ₂	0.78	0.1557	4	0.74	0.1643	1.49
RN_{LC}	A ₆ , B ₃ ; C ₂	0.77	0.2956	2	0.70	0.2540	1.41
RN_{SG}	A ₀ , B ₁ ; C ₀	0.45	0.0107	4	0.26	0.0097	0.99
RN_{SS}	A ₆ , B ₁ ; C ₀	0.46	0.3755	4	0.36	0.3234	1.07
RN_{SU}	A ₅ , B ₁ ; C ₂	0.64	0.1404	6	0.57	0.1533	1.21

^a See table 3.1 footnote.

Table 3.4. PLSR results for predicting some potato constituents using VIS/NIR hyperspectral imaging (whole tubers) for Frito Lay 1879 and Russet Norkotah cultivars.

Cultivar _{Constituent}	Preprocessing ^a	Calibration			Prediction		
		R _{cal}	RMSEC _{CV}	LVs	R _{pred}	RMSEP	RPD
FL_{GL}	A ₉ , B ₁ ; C ₀	0.77	0.0770	4	0.38	0.0681	0.93
FL_{LC}	A ₆ , B ₁ ; C ₀	0.49	13.124	7	0.47	11.7014	1.14
FL_{SG}	A ₀ , B ₁ ; C ₀	0.22	0.0112	2	0.19	0.0148	1.02
FL_{SS}	A ₅ , B ₁ ; C ₀	0.34	0.4629	2	0.24	0.4602	1.03
FL_{SU}	A ₉ , B ₁ ; C ₀	0.18	0.0817	1	0.14	0.0702	1.02
RN_{GL}	A ₄ , B ₁ ; C ₀	0.75	0.3669	4	0.52	0.3259	1.19
RN_{LC}	A ₇ , B ₁ ; C ₀	0.78	9.5766	5	0.43	7.8047	1.10
RN_{SG}	A ₄ , B ₁ ; C ₀	0.30	0.0107	2	0.20	0.0095	1.01
RN_{SS}	A ₅ , B ₃ ; C ₀	0.55	0.4242	9	0.29	0.4277	0.81
RN_{SU}	A ₅ , B ₃ ; C ₀	0.44	0.1879	2	0.43	0.1805	1.10

^a See table 3.1 footnote.

3.3.3.3 Results for transmittance mode

Taking into account both R and RPD values, the transmittance mode yielded strong correlations for leaf counts in the case of FL with R and RPD values of 0.87 and 1.94 and for RN the values were 0.81 and 1.54 respectively (Table 3.5). The glucose prediction model for RN also showed close correlation performance to the interactance mode with sliced samples with R

and RPD values of 0.87 and 2.01, but lower correlation was obtained for FL with R and RPD values of 0.66 and 1.23. The sucrose model for RN yielded comparable results to those obtained using interactance mode with sliced samples with the values of R and RPD as of 0.63 and 1.30 and for FL, the values were 0.57 and 1.23. Prediction models for specific gravity and soluble solids didn't show as high correlation performance as other three constituents.

Table 3.5. PLSR results for predicting some potato constituents using NIR transmittance (sliced samples) for Frito Lay 1879 and Russet Norkotah cultivars.

Cultivar _{Constituent}	Preprocessing ^a	Calibration			Prediction		
		R _{cal}	RMSEC _{CV}	LVs	R _{pred}	RMSEP	RPD
FL _{GL}	A ₀ , B ₁ ; C ₁	0.90	0.0750	9	0.66	0.0515	1.23
FL _{LC}	A ₆ B ₁ ; C ₂	0.97	0.2788	20	0.87	0.2587	1.94
FL _{SG}	A ₀ , B ₃ C ₀	0.66	0.0033	1	0.56	0.0036	1.22
FL _{SS}	A ₇ , B ₃ C ₀	0.40	0.5335	1	0.30	0.4509	1.05
FL _{SU}	A ₅ , B ₁ ; C ₀	0.60	0.0782	10	0.57	0.0582	1.23
RN _{GL}	A ₁₂ , B ₁ C ₀	0.96	0.2319	9	0.87	0.1921	2.01
RN _{LC}	A ₅ , B ₁ C ₁	0.90	0.3383	13	0.81	0.3453	1.54
RN _{SG}	A ₇ , B ₁ ; C ₀	0.69	0.0101	6	0.59	0.0079	1.22
RN _{SS}	A ₄ , B ₃ ; C ₀	0.87	0.6281	3	0.23	0.5938	0.58
RN _{SU}	A ₅ , B ₁ C ₁	0.73	0.8555	10	1.07	0.63	1.30

^a See table 3.1 footnote.

3.4 Discussion

The results indicate three modes (interactance, transmittance and hyperspectral) used to build prediction models for some constituents in potato tubers have dependable results for leaf primordium leaf counts (comparable to the work conducted by Jeong et al. (2008), and glucose and sucrose (comparable to the work conducted by Mehrubeoglu and Cote (1997); and Hartman and Buning-Pfaue (1998)). The transmittance mode was inferior in performance for these three constituents. A note to make is that for the interactance mode, the whole tubers yielded similar performance for the prediction models of leaf count and glucose for FL compared with the sliced samples which is important as it could save processing time in terms of measurement and sampling for commercial application, and is nondestructive. In general, specific gravity, which is strongly related to dry matter, and soluble solids were not well predicted using the systems and models presented here which contrasts with some other research in the literature (Hartman and

Buning-Pfaue, 1998; Haase, 2004; Dull et al., 1989; Subedi and Walsh, 2009; Chen et al., 2005; Scanlon et al., 1999). This study presents the application of spectroscopic and hyperspectral imaging technologies, plus modeling, toward addressing a significant issue of rapid detection of reducing sugars, that are very critical to the frying industry, which does not currently exist in the market for the purpose of quality management and potato industry profitability.

3.5 Conclusions

NIR transmittance in the range of 900-1685 nm, visible/near infrared interactance spectroscopy in the range of 503-1047 nm, and hyperspectral reflectance, in the range of 400-1000 nm, were used to build prediction models to measure constituents in potato tubers that are important to chipping and seed potato industries. Two cultivars were used to conduct the study, FL and RN. The study showed that the prediction of leaf count and glucose, and somewhat lesser for sucrose, was possible using interactance, in both sliced samples and whole tubers, and in less degree using hyperspectral reflectance and transmittance systems, for sliced samples, for FL. However, interactance and transmittance, for sliced samples in both modes, showed possible reliable prediction for RN. It is worth to note that both cultivars showed strong correlation for the sliced samples and the whole tubers only in the case of interactance mode. Specific gravity and soluble solids prediction models are weak and further improvement is necessary to obtain reliable models. Thus, while previous studies of the application of visible/NIR techniques to estimate sugars demonstrated good results, it should be noted that they were conducted on homogenized samples (Hartmann and Buning-Pfaue, 1998) or without validation on different sets of data (Mehrubeoglu and Cote, 1997). Also, results for leaf counts prediction using NIR conducted by Jeong et al., (2008) did not include confirmation. This study included validation

data sets and measurements of intact potato tubers or slices thus leading to more confident results and more direct practical industry applicability.

CHAPTER 4 EVALUATION OF SUGAR CONTENT OF POTATOES USING HYPERSPECTRAL IMAGING SYSTEMS

(Rady, A.M., Guyer, D.E., Lu, R. 2014. Evaluation of sugar content of potatoes using hyperspectral imaging. Journal of Food Bioprocess and Technology (in review and initially accepted))

4.1 Introduction

Hyperspectral imaging (HI) for agricultural applications has been studied for two decades. The technique requires minimal sample preparation and is non-destructive with the capability of visualizing the spatial distribution of desirable constituents (El Masry and Sun, 2010a). It was used for detection of defects and surface contaminants and estimation of internal constituents in food (Lawrence et al., 2001; Qin and Lu, 2007; Molto et al., 2010; El Masry and Sun, 2010b; Chao, 2010; Menesatti et al. 2010; Wang and El Marsy, 2010). Qiao et al. (2005), studied hyperspectral imaging to estimate both the water content and the weight of potato tubers. The system was used to extract morphological features and spectral responses to the water content in tubers simultaneously. The wavelength range of 934-997 nm was found to be useful for predicting the water content in potato tubers. Results showed that the coefficient of correlation between the predicted and actual values of water content was 0.93 and 0.77 for training and validation, respectively. Water content is an important factor for potato tubers as it is positively proportional to the yield and consequently the total profit for the grower. Singh et al. (2004), developed a partial least squares model, using a spectroradiometer, for prediction of the potato tuber water content with the correlation coefficient being as high as 0.99.

Lu and Peng (2006), developed a hyperspectral imaging-based spectral scattering technique to estimate peach firmness. A Lorentzian distribution function was used to model the

scattering profiles and multi-linear regression (MLR) was then applied to build the calibration model, which resulted in coefficients of determination (R^2) of 0.67 to 0.77 for the validation data. Mehl et al. (2002), developed a hyperspectral imaging system (HIS) for detection of various apple defects, including bruises and diseases. They selected three best wave bands for classification of apple defects with the classification rates being 100%, 63%, and 70% for ‘Gala’, ‘Delicious’, and ‘Golden Delicious’, respectively, for the normal samples, and 100%, 63%, and 68% for the defected samples. The advantages of HI as an accurate technique of non-destructive defect evaluation of food products and more importantly obtaining few wavelengths that are strongly associated with high classification rate, encourages the application of HI in constituents evaluations. However, studying HI systems in constituent prediction alone is insufficient.

Consequently, the combination of constituent-sorting with the traditional damage-based sorting can be more reliable, cost and time effective and robust than using multi-stage sorting systems or combining vision and spectroscopic systems together to achieve the goal of monitoring tuber quality from different perspectives. The objectives of this study were:

1. Determine the potential of hyperspectral imaging systems for quantifying the levels of sucrose and glucose in potato sliced samples for two different-use cultivars.
2. Develop prediction models for estimating the amount of sucrose and glucose in potato tubers covering levels used to assess suitability of tubers for processing which are important for potato growers and processors.
3. Develop classification models for potato tubers of both Frito Lay 1879 and Russet Burbank based on sugar levels and using multiple methods.

4.2 Materials and Methods

4.2.1 Raw Material and Experimental Design

Experiments were conducted in 2009 and 2011, and in both seasons, two common cultivars were used in the experiments, Frito Lay 1879 (FL) which is a chipping cultivar, and Russet Norkotah (RN) which is usually used fresh for baking and boiling. The experimental setup and design for the 2009 and 2011 seasons are shown in Fig. 4.1 and Fig. 4.2 respectively. In the 2009 season, the RN cultivar was hand-harvested from a research farm at Montcalm, MI. (sandy soil). There were two vine killing treatments (0 and 7 days from Aug. 13th), with each followed by three harvesting periods (7, 14, and 21 days following the vine killing). The FL cultivar was harvested from two different farms: the Montcalm research farm, in which there were two vine killing dates each followed by three harvesting dates as with RN, and the MSU Muck experimental farm (muck soil), Bath, MI in which there were six vine killing treatments (0, 7, 14, 21, 28, and 35 days from Aug. 13th) followed by three harvesting periods for each vine kill. Samples were stored in three temperatures of 4, 7, and 10°C. Tubers were then monthly sampled for experimentation starting in November, 2009 until April, 2010 (except at March) with a total number of 540 tubers from FL and 180 tubers from RN tested through the 2009 experiments. The sampling procedure was designed to obtain a broad range of sugar content samples.

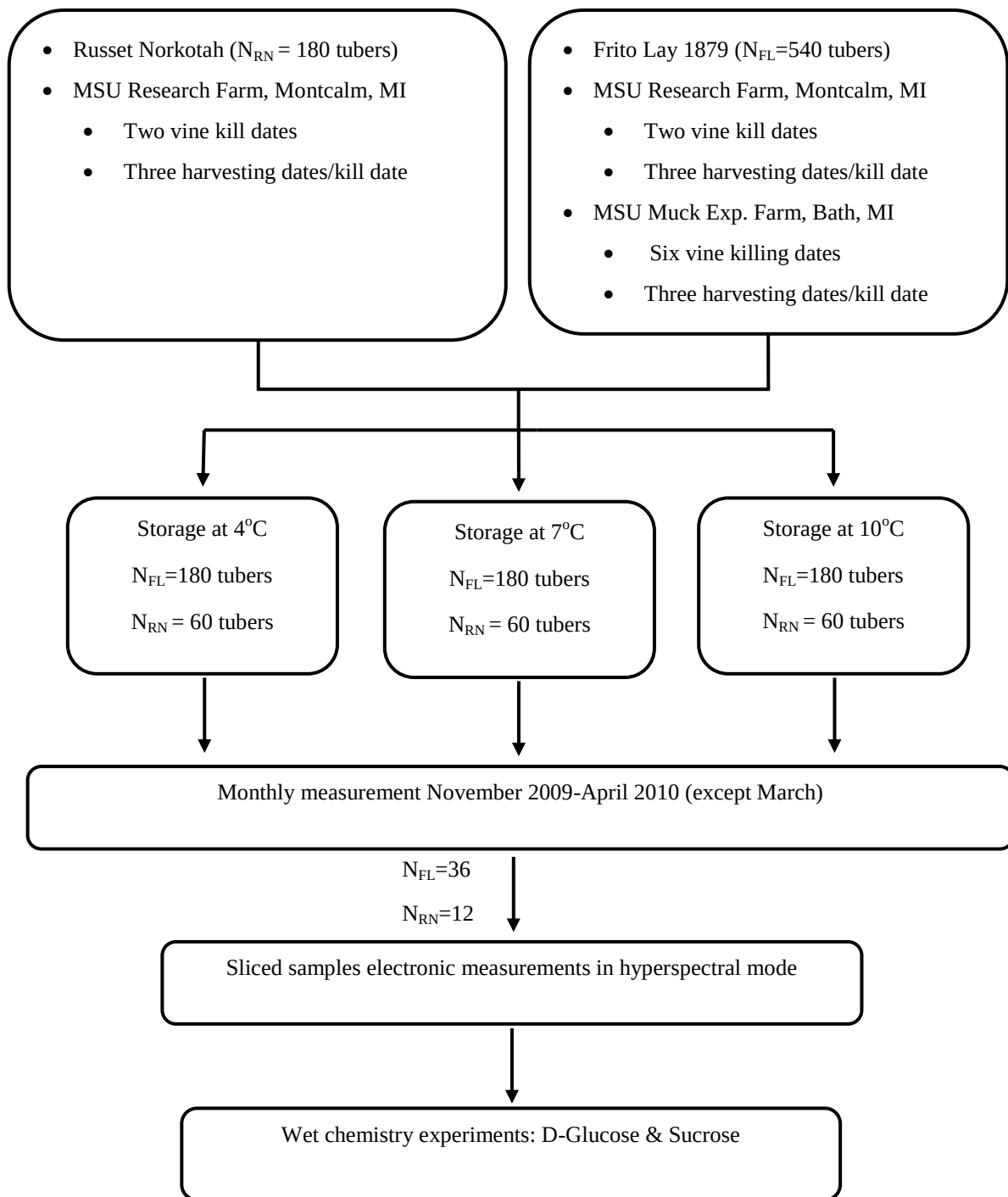


Figure 4.1. Flow chart of the experimental design to assess physiological status of potato tubers using VIS/NIR hyperspectral imaging for Frito Lay 1879 and Russet Norkotah cultivars in the 2009 season.

In the 2011 season, both cultivars were obtained from a commercial production field (sandy soil) in Southwest Michigan. Samples were hand-harvested on only one date in September, 2011. Two more storage temperatures were added in order to obtain more uniform sugar distribution and simulate the various uses of potato tubers. In general, lower storage temperature is desired for cultivars that are used as seeds or for cooking, while higher temperatures are used for chip cultivars. Tubers were first stored at 4 °C for three weeks and an initial electronic measurement was conducted. Tubers were then distributed over five different cold storage rooms with the following temperatures: 1, 4, 7, 10, and 13 °C. They were then sampled for experimentation starting in November 2011, and each month until May 2012 (except at April) with a total number of 195 tubers from FL, and 75 tubers from RN. In both seasons, tubers were cleaned prior to the imaging, and any defective samples were discarded.

It is important to emphasize that the main target of collecting samples from different locations and storing at different temperatures was to obtain broad, and uniform, sugar distribution, rather than evaluating the growing condition, and other pre- and post-harvest practices that were conducted on tubers. Consequently, results representing different locations for Frito Lay1879 were not separately analyzed and compared.

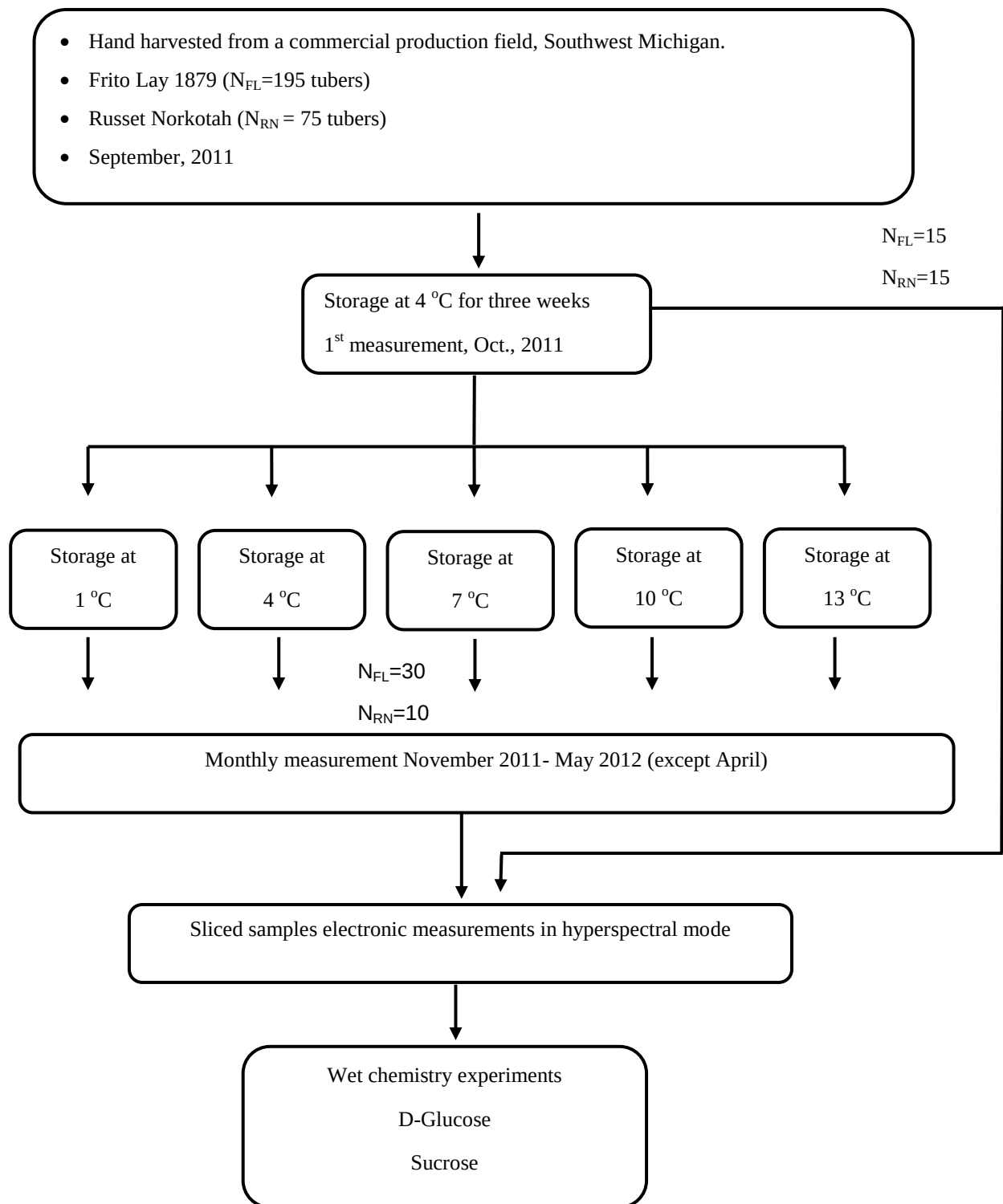


Figure 4.2. Flow chart of the experimental design to assess physiological status of potato tubers using VIS/NIR hyperspectral imaging for Frito Lay 1879 and Russet Norkotah cultivars in the 2011 season.

4.2.2. Constituent Measurement

4.2.2.1 Potato sample preparation

Sample slices were used for VIS/NIR hyperspectral reflectance imaging. Each slice was 0.5 inch (12.7 mm) thick and it was obtained by cutting the tuber three times in a direction that is perpendicular to its longitudinal axis, starting from the stem end of the tuber. The tested slice was the third slice in the cutting routine.

4.2.2.2 Wet chemistry basis measurements

This destructive process was to provide a basis for the amount of sucrose and glucose in the tuber or piece of tuber that has been subjected to the electronic measurements. The standard method used to estimate glucose and sucrose is the enzymatic method, using the glucose oxidase and paraoxidase enzymes.

4.2.2.2.1 Extraction of juice

Immediately after the electronic measurement, each slice was put in a plastic bag and stored in a foam box contacting ice to maintain the sample in a fresh state and minimize any chemical changes during the period of performing electronic measurements for other samples. To ensure consistency between the slice electronic and wet chemistry measurements, a sufficient amount of potato tuber juice from the specific areas that had already been electronically tested was obtained by using a 1 inch (25.4 mm) cylindrical metal core borer to extract tissue primarily from the middle of the slice. This tissue was then put in a pre-sterilized 7 oz Whirl-Pak filter bag, 9.5 x 18 cm (Nasco, Fort Atkinson, Wisconsin, USA). The bag was then hammered by hand using a 2 lb weight for juicing and then homogenized using a stomacher for 1 min. The juice was filtered by the Whirl-Pak filter bag and transferred with a pipette to a polystyrene tube with cap.

This juice was stored at -20 °C to reduce any variation of constituents and allow subsequent use and analysis of the juice at a later time.

4.2.2.2 Performing the chemical estimation of glucose and sucrose

Using the Megazyme sucrose/D-glucose assay procedure (Megazyme International Ireland Ltd), the ratio of each of glucose and sucrose, gram per gram fresh tuber weight, was measured and calculated using the same approach noted in section 3.2.3.1.2.

4.2.3 VIS/NIR hyperspectral imaging systems

Two hyperspectral imaging systems were used in this project and both detected the diffuse reflected light from the sliced samples. The first system used in the 2009 season was the same as noted and described in section 3.2.2.3.

In 2011, a different hyperspectral imaging system was used for the experiment, because the system used in 2009 was no longer available for the research. Although the two systems were quite similar in measurement principle, they were dissimilar enough that the models of both seasons were separated and no combining of data was conducted. The 2011 system, the Optical Properties Analyzer or OPA, was developed at the postharvest engineering lab of USDA-ARS (Cen, and Lu, 2009).

The OPA system consists of: a high performance 14-bit electron-multiplying CCD camera (Luca^{EM} R604, ANDORTM Technology, South Windsor, Connecticut, USA) covering the wavelengths of 400-1000 nm; a monochrome megapixel frame transfer sensor with 1004x1002 pixels of 8x8 µm, thermoelectrically cooled to -20°C; and an enhanced imaging spectrograph (ImSpector V10E, Spectral Imaging Ltd., Oulu, Finland) directly connected to the CCD camera. Point scan mode was used in the experiments and was conducted using a prime

lens (Xenoplan 1.9/35-0901, Schneider Optics, Hauppauge, NY, USA). The light source used was a tungsten halogen light bulb with 20 W output power (HL-2000-HP, Ocean Optics, Dunedin, FL, USA) connected to a DC regulated controller chip (PT6201N, 12, Texas Instruments Inc., Dallas, Texas, USA) to provide point light. The light beam at the focal point was 1 mm diameter provided by an optical fiber coupled with a focusing lens. The incident light is 1.6 mm away from the scanning line and is 15° to the vertical axis (Cen and Lu, 2009). During the scanning process, 11 images were acquired along a movement distance of 5 mm of the horizontal stage with a resultant image size of 251x 250 pixels with a spatial resolution of 0.21 mm/pixel. A close view of the system clarifying the sample holder that slides horizontally using the stepping motor is shown in Fig. 4.3a, and a schematic configuration of the 2011 hyperspectral system is shown in Fig. 4.3b.

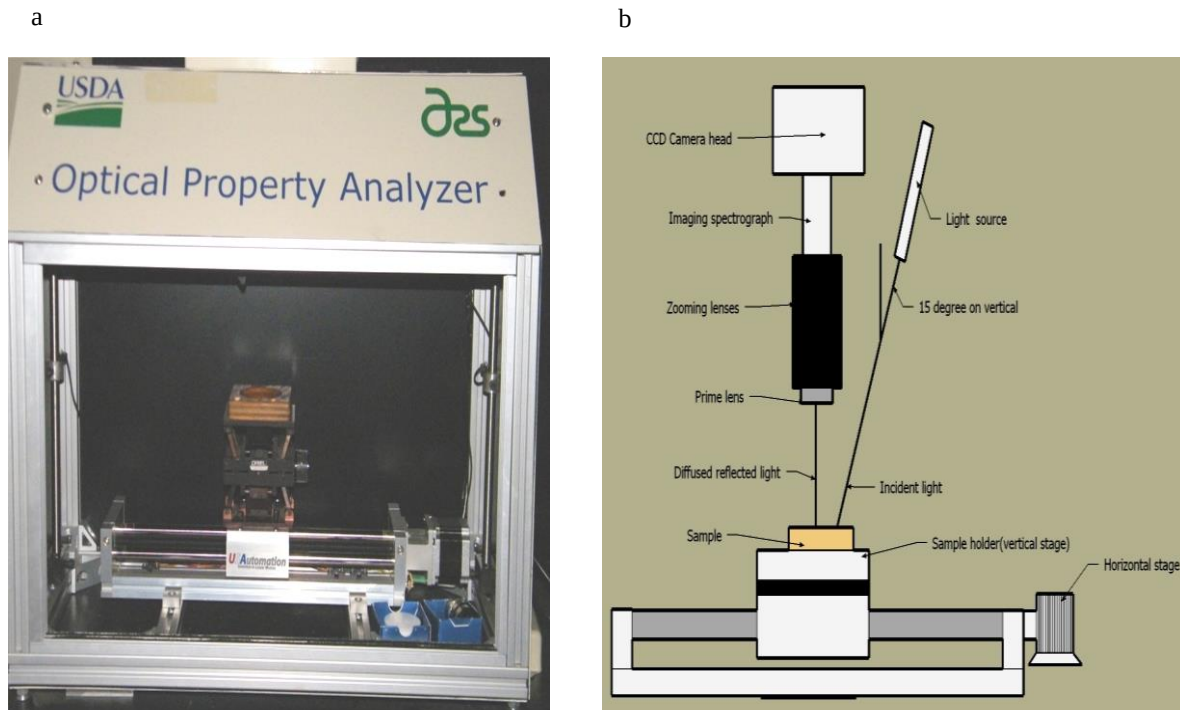


Figure 4.3, a. Hyperspectral imaging Optical Properties Analyzer (OPA) used in the 2011 season. b, Schematic of OPA.

The light scattering behavior inside the sample subjected to incident light is shown in Fig. 3.3.b. The light radiation beam penetrated the sample surface and scattered outward through the tissue, and the backscattered light was captured by the hyperspectral imaging system in line scanning mode. The primary difference between the 2009 and 2011 measurements was that the light radiation beam had a diameter of 1.5 mm at the focal point for the 2009 season and 1 mm for 2011 season. The raw output of both systems was the same as in Fig. 3.4. It should be noted that both sides of each slice were tested and consequently the total number of samples was 1080 for FL and 360 for RN in the 2009 season. In the 2011 season, there were 390, and 150 sliced samples for FL, and RN respectively.

4.2.4 Data Analysis Discussion and Approach

In this section, feature extraction, and methods of building calibration and prediction models are explained in detail. Several analysis methods were also added in this chapter in comparison to chapter 3. In addition to mean reflectance spectra, curve fitting parameters were also extracted using an exponential model. Several types of artificial neural network were used to build training and testing models for sugar prediction. Moreover, wavelength selection techniques (interval partial least squares and genetic algorithm) were also added to detect the most influencing variables associated with yielding strong correlation between optical measurements and sugar concentrations. Finally, classification of potato tubers of both FL and RN based on sugar levels was conducted using multiple common classification techniques.

4.2.4.1 Definition and development of descriptive variables

4.2.4.1.1 Extracted mean spectra

The average reflectance spectra for the hyperspectral data were obtained for the sliced samples in the case of both cultivars. To extract the average reflectance for each image, all

wavelengths from 400.9 to 1000.1 nm were considered as shown in Fig. 4.4a. At each wavelength, the arithmetic mean of intensity values of the spectra, as shown in Fig. 4.4b, was calculated. Finally, a 1*256 vector array for the 2009 season as shown in Fig. 4.4c or 1*250 for the 2011 season is obtained from each image. The same process is repeated for each of the 10 images per sample and the average is calculated to represent one sliced sample. All mean reflectance spectra is divided by the equivalent spectra of standard Teflon[®] resulting in a relative mean reflectance spectrum for each sample.

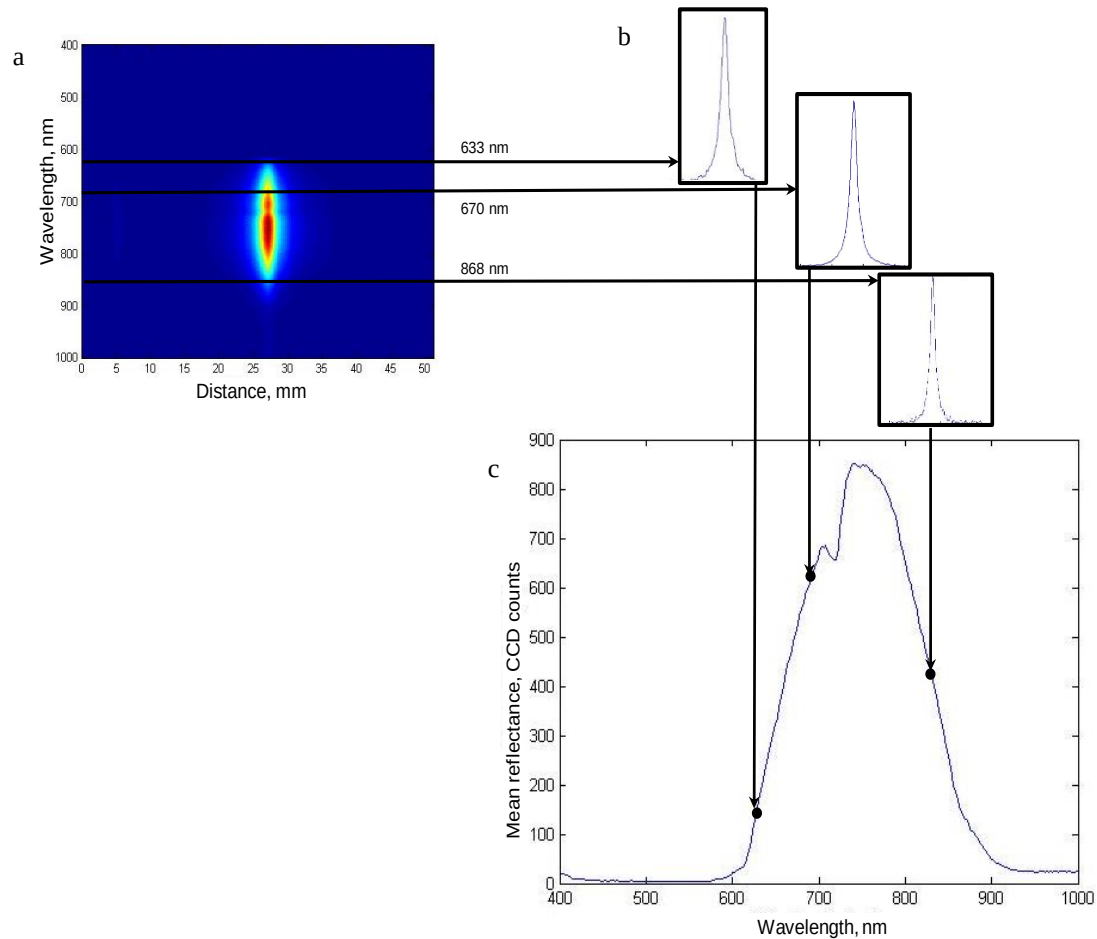


Figure 4.4, a. An example of an image obtained for each slice sample, b. Sample of spectra at different wavelengths, c. Sample of average spectrum for one image.

4.2.4.1.2 Describing scattering profiles

In addition to the mean reflectance spectra extracted data (section 4.2.4.1.1), describing or fitting scattering profiles (or original reflectance curves) was conducted on the relative reflectance curves to obtain more information about sample behavior under the studied hyperspectral systems. The approach of modeling scattering profiles was successfully applied on apple (Peng and Lu, 2005; Peng and Lu, 2007a; Peng and Lu, 2004; Peng and Lu, 2007b) and on peach (Lu and Peng, 2006). In such approach, the scattering profile is described using Lorentzian distribution, exponential distribution, or Gaussian distribution with different numbers of parameters for each model. In the current study, all three distributions were applied. A preliminary, exponential distribution with two parameters was found to be the best model to simulate scattering profiles for potato slice samples in the 2009 and 2011 seasons with the following equation describing the exponential model:

$$I_w = a_{w_i} e^{\left| -\frac{x}{b_{w_i}} \right|} \quad (4.1)$$

Where I_w is the light intensity at wavelength w_i in CCD counts; x is the scattering distance measured from the beam (mm); a_{w_i} represents the intensity peak value in CCD counts for the scattering profile when $x=0$; and b_{w_i} is the scattering width, in mm, at half (0.37) of the intensity peak value; and the subscript w_i is the wavelength in the range 400-1000 nm with $i = 1, 2, \dots, n$ where n is the total wavelengths used. Both sides of each spectral profile were averaged before conducting the curve fitting. Scattering profiles used in curve fitting parameters were covering a spatial scattering distance of 8 mm (or 42 pixels) for 2009 season and 4 mm (or 22 pixels) for 2011 season. Choosing both distances was to avoid using noisy areas that might affect the accuracy of calculating curve fitting parameters. To estimate a_{w_i} , and b_{w_i} , a nonlinear

regression technique was applied for each scattering profile in the considered wavelength range for each season using the curve fitting tool box in Matlab[®] software (version 7.5.0.342, MathWorks, Natick, MA, USA).

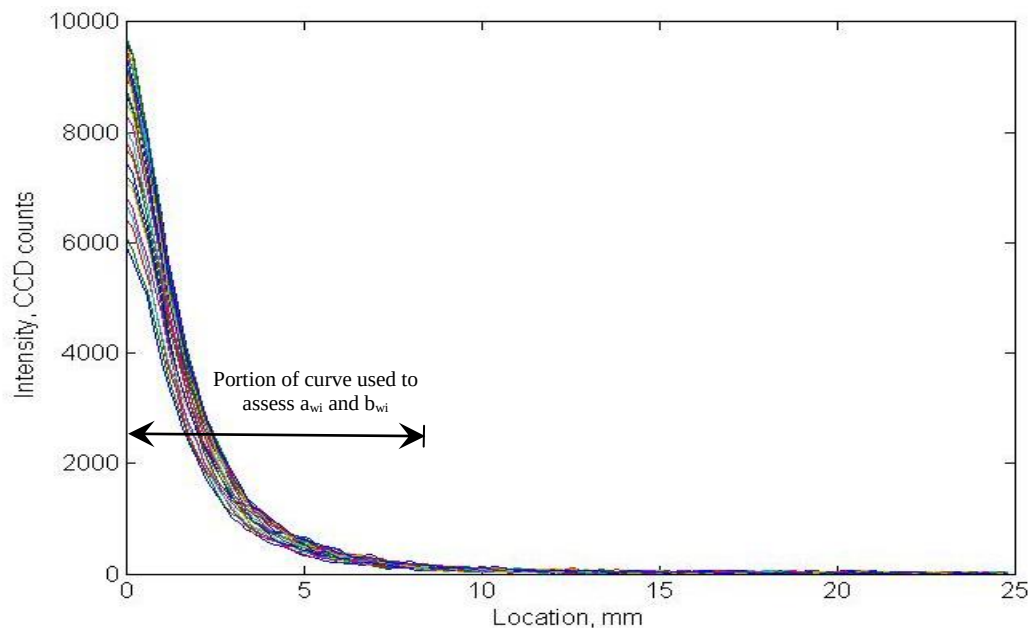


Figure 4.5. Decaying portion of original spatial scattering profiles for selected sliced samples of Frito Lay 1879 cultivar at 698.7 nm in the 2009 season.

4.2.4.2 Partial least squares regression (PLSR)

A complete description of PLSR used in this research along with pretreatment for either spectra or reference values is covered in section 3.2.4. It should be noted that three types of data sets were used: only the mean spectra, the two curve fitting parameters (a_{wi} , b_{wi}) concatenated to each other, and finally combining all mean and curve fitting parameters.

It is worth stating that according to William (2007), correlation coefficient (R) value was used to evaluate prediction model efficacy. Values of R of 0.81-0.90 can be used for screening and approximate calibration. Whereas, R values of 0.91-0.95 may be carefully used for most applications. The prediction models with R values above 0.95 are appropriate for quality

assurance. RPD values of 1.5-2.0 are capable to differentiate between high and low constituent values, while values of RPD in the range of 2.0-2.5 means a possibility of coarse prediction of reference values. Values of RPD of 2.5-3.0 or higher can be used for good and excellence prediction, respectively (Nicolai et al., 2007).

4.2.4.3 Artificial neural network (ANN)

ANN, which are broadly used in classification tasks, are computational algorithms that may be used to gain an understanding of biological systems. An artificial neural network is a machine that is designed to mimic the method of that of the brain when it conducts a certain task (Haykin, 2009). From the regression side, PLSR is a technique that depends on building calibration models using linear combination of independent variables and other coefficients that are determined during a training (or calibration) process. ANN, however, depend on training the data in a non-linear mapping from the independent variables into another stage or layer (called hidden layer) followed by a linear mapping from the hidden space to the output space that just contains the reference value (glucose or sucrose concentration). Two types of ANN were used to obtain prediction models for each constituent of interest: the radial basis functions neural networks (RBFNN) and the feed forward neural network (FFNN). The RBFNN consisting of choosing a function $F(x_i)$ that satisfies the following constrain:

$$F(x_i) = y \text{ for } i = 1, 2, 3, \dots, n$$

Where n refers to sample size, x refers to a vector of independent variables (wavelengths). In RBFNN, $F(x_i)$ is chosen as follows:

$$F(x) = \sum_{i=1}^n \omega_i \phi (\|x - x_i\|) \quad (4.2)$$

Where ω is a weight vector, $\phi(\|x - x_i\|)$ is a set of nonlinear functions known as radial basis functions, $\|\cdot\|$ denotes a norm that's the Euclidean distance, and x_i is a point located in the center of the radial basis function. Equation 4.2 can be rewritten in the matrix form as follow:

$$\Phi W = y \quad (4.3)$$

Where $\Phi = \{\phi_{ij}\}_{i,j=1}^n$ is an N by N matrix with elements ϕ_{ij} , W an N by 1 vector containing weights, and y is N by 1 vector containing reference values. Then W can be found as Φ is a non-singular matrix. RBFFNN consists, as shown in Fig. 4.6, of the following layers:

1. Input layer: consists of m variables each representing one of the extracted features (mean reflectance, concatenated a_{wi} , and b_{wi} , and concatenated mean spectra, a_{wi} , and b_{wi}).
2. Hidden layer: consists of a certain number of neurons, the radial basis functions were in this case chosen as a Gaussian function as follow:

$$\phi_i(x) = \phi(\|x - x_i\|) = e^{-\frac{1}{2\sigma^2}\|x - x_i\|^2}, \quad i = 1, 2, 3, \dots, n \quad (4.4)$$

Where σ is the spread or width that was chosen as 3. The number of neurons was chosen as 1000 units.

3. Output layer: This represents the predictor variable that is in fact the glucose or sucrose concentration. The allowable mean square of error (MSE) was selected as 0.0001.

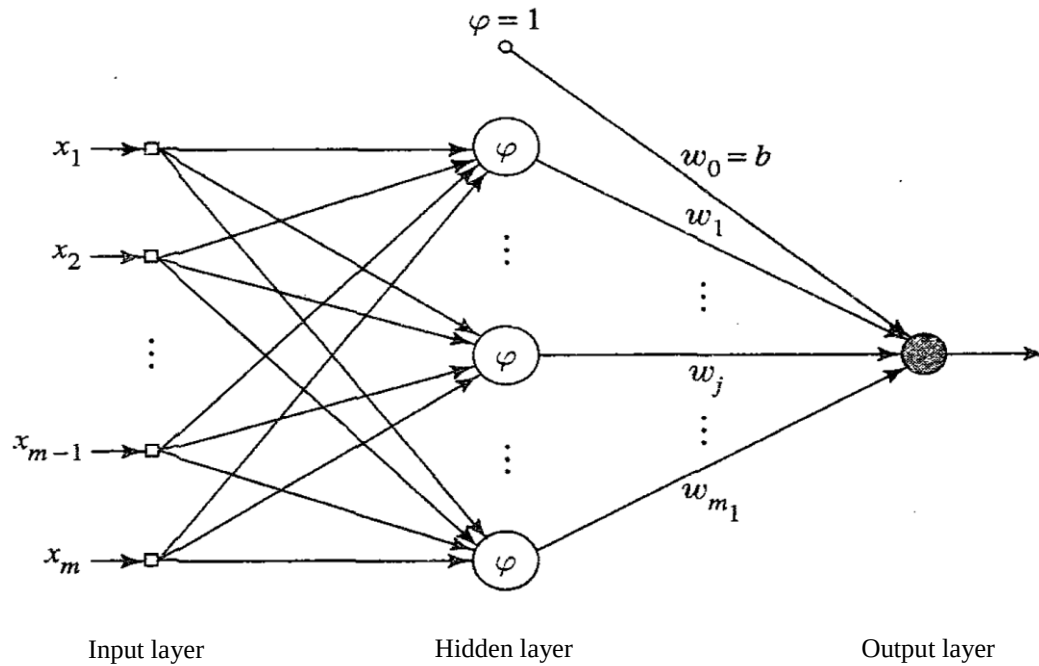


Figure 4.6. Schematic representation of RBFNN (after Haykin, 2009).

Another sub type is an exact design RBFNN (RBFNNE) in which the network is able to produce a zero-error training vector containing as many neurons in the hidden layer as the number of independent variables (wavelengths). The spread was chosen as in the regular RBFNN, 3.

The last type of the radial basis function neural networks is the generalized RBNN (NEWGRN) contains four layers. The first layer contains the input values (mean reflectance, concatenated a_{wi} , and b_{wi} , and concatenated mean spectra, a_{wi} , and b_{wi}), the second layer is a hidden layer including as many neurons as the number of wavelengths. The third layer also includes as many neurons as the number of wavelengths but with different bias weight set to the target (sugar concentration). The final layer contains glucose or sucrose concentration.

The data was randomly divided into 75% for training the network, and 25% for testing the network and creating the prediction model. The training set was then divided into four sub sets, and then a four-fold cross validation technique was used to obtain the best training model

based on the root mean square of error of cross validation of the training set ($SeCV_{train}$). The predicted reference values were obtained from the testing spectral data when substituted into the best trained model and then they were compared with the actual reference values, both correlation coefficient (R_{test}) and root mean square error (Se_{test}) were then calculated.

The second type of artificial neural network used in obtaining the prediction models was feed forward with back propagation network (FFNN) also known as multilayer perceptron. In this network type, first, N linear combinations of the x-variables (spectra) are built as in the following equation:

$$v_j = a_{0j} + a_1x_1 + a_2x_2 + a_3x_3 + \dots + a_mx_m \text{ for } j = 1, \dots, N \quad (4.5)$$

And then a nonlinear function, called the activation function, usually a sigmoid type is applied as follows:

$$z_j = f(v_j) = \frac{1}{1+\exp(-v_j)}, \text{ for } j = 1, \dots, N \quad (4.6)$$

Finally, the predicted output, $\hat{y}$, is calculated as a linear combination of the values from different neurons as follows:

$$\hat{y} = b_0 + b_1z_1 + b_2z_2 + b_3z_3 + \dots + b_Nz_N$$

Where a_0 , b_0 are called bias and assumed to be equal to 1. $a_1, a_2, \dots, a_N$, and $b_1, b_2, \dots, b_N$ are weights determined during the training process. The back propagation algorithm is a common technique in training FFNN and it's an extension of least mean squares algorithm and is based on gradient descent in error and consequently weights updating.

In this research, FFNN consists of an input layer which represents the mean relative reflectance spectra, concatenated a_{wi} , and b_{wi} , or concatenated mean spectra, a_{wi} , and b_{wi} for each sample. The number of neurons in the hidden layer contained several trials including 50, 100,

150, 200, 250 and 300 neurons, and the output layer contained one neuron which is the real value of reference (constituent). The transfer function for the hidden layer is the tan-sigmoid function, and for the output layer is a linear transfer function. The training style was chosen to be the scaled conjugate gradient method. A schematic view of the FFNN is shown in Fig. 4.7.

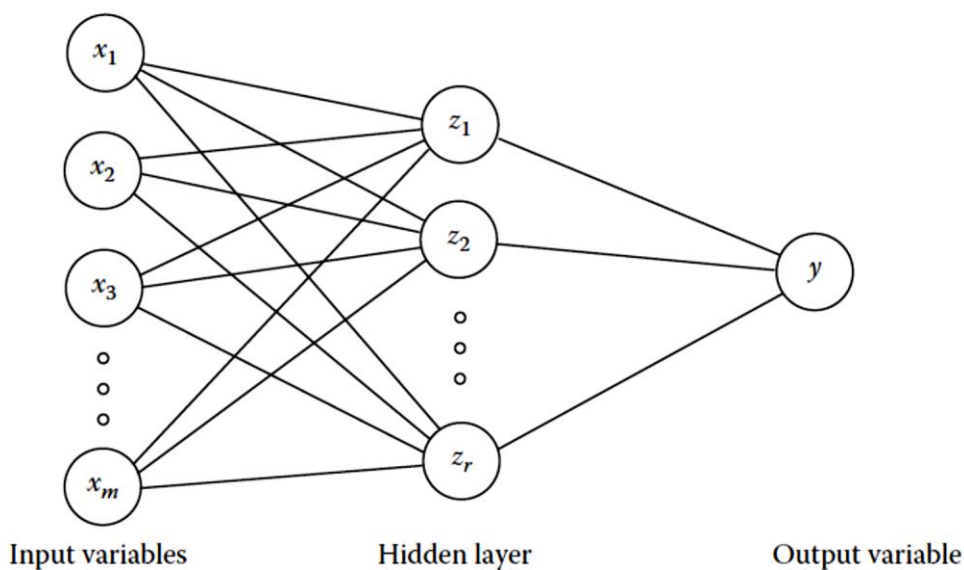


Figure 4.7. Schematic representation of FFNN (after Varmuza and Flizmoser, 2007).

In the FFNN method used in this research, the data was randomly divided into three groups: the first one is used to train the network and it was around 60% of the samples. The second group represented about 20% of the samples and it was used to validate the built network and four-fold cross validation technique was used to obtain the best calibration model. The third group is an independent set to test the network. The predicted values of reference were compared with the actual values and then both the correlation coefficient and the root mean square error were calculated. The stopping rule in this case is when the mean square error (mse) of the validation set of data reaches a minimum or sequential number of iterations is reached.

4.2.4.4. Wavelength selection

Variable selection techniques help identify subsets of variables (wavelengths) for a given problem which yield the most powerful and accurate model. In multivariate analysis, using all variables may produce a better fit for building calibration models as a higher number of variables may yield smaller residuals and consequently a better R value. However, the more important goal is to optimize the prediction model performance for the validation set of data. Reducing the number of regressors can overcome potential problems of overfitting (Varmuza and Filzmoser, 2008). Moreover, measuring certain variables can be difficult and/or other variables may contain noise or signals which interfere with the signals which are valuable for compound detection. Two methods of variable (wavelength) selection were used in this research; the interval partial least squares (IPLS) and genetic algorithm (GA). Configurations of both IPLS and GA were based on preliminary analysis that led to using the following parameters based on the performance of PLSR and ANN prediction models. The IPLS method is a known variable selection method for spectroscopic data and for optimizing the performance of PLSR models. IPLS uses sequential and exhaustive methods of search for the best subset of variables in either a forward or a backward direction and different window width values (number of variables per window). In this research, forward mode, windows of 1, 2, and 3 variables, with number of latent variable for the PLSR model being 15, were used.

The genetic algorithm mainly depends on randomly selecting different subsets of variables called chromosomes or individuals and in each chromosome some variables (genes) are selected or active, denoted by 1, and others are not selected, denoted by 0. With the use of cross validation, each individual prediction model will have its fitness (commonly root mean square error of cross validation ($RMSE_{CV}$)). Based on the fitness threshold, some chromosomes are

discarded and others, the remaining individuals, are used to build new chromosomes by using crossover and mutation methods. Finally, the process of evaluation and forming new chromosomes is repeated until a highest fitness, i.e. lowest $RMSE_{CV}$, chromosome is obtained. In the current study, window width values of 1, 2, 3, double crossover, maximum number of generations of 300, maximum number of partial least squares latent variables of 20, and three iterations were used in the forming of genetic algorithm. After reviewing PLSR and ANN prediction models for both seasons in the case of FL and RN cultivars and observing that there was close prediction results between mean, curve fitting parameters, and combined mean and curve fitting parameters, only mean reflectance spectra were used for variable selection.

4.2.4.5 Classification of potatoes based on sugar levels

Sorting tubers based on sugar levels was conducted using two common techniques K nearest neighbor (Knn) and partial least squares discriminant analysis (PLSDA). As a nonparametric classification method, Knn requires no model to fit or classify the point (sample). However, the distance, usually Euclidean, between the point and the selected neighbors (k) is calculated. The sample is then classified to the nearest class or to the class having the majority vote (Varmuza, and Filzmoser, 2007; Wise et al., 2006; Bishop, 2006; Duda et al., 2001). In this study, the k values were selected as 3, and 5. PLSDA is a linear regression classification-based method that is similar to linear discriminant analysis (LDA) with the advantage of noise reduction and latent variable selection being in PLSDA (Wise et al., 2006). In this study, spectra data and reference variables were preprocessed as mentioned in section 2.4.2, with 10-fold cross validation used to increase the robustness of the training models for both methods. Samples were divided into two classes based on sugar values with cut-off values of glucose for FL and RN as of 0.035 and 0.035% respectively, whereas the values for sucrose were 0.03 and 0.10%. Cut-off

levels were adopted from recommended thresholds listed by Stark and Love, (2003), for both sugars except for the sucrose level for FL which was chosen to create two balanced classes. Only mean reflectance spectra (MRS) data was used for samples classification with the note that selected wavelengths using IPLS were applied to MRS data and the results were used in classification tasks. Data was divided into training (75%) and testing sets (25%). Classification of sugars was conducted using the classification toolbox for Matlab created by Davide Ballabio (Milano Chemometrics and QSAR Research Group, University of Milano - Bicocca, Milan - Italy) and the PLS routine used to compute PLS-DA was written by Frans W.J. van den Berg (Quality & Technology group, section Spectroscopy and Chemometrics, Department of Food Science, University of Copenhagen).

4.3. Results and Discussions

4.3.1 Distribution of Glucose and Sucrose

Table 4.1 shows the statistics of glucose and sucrose for all samples (reference variables) based on wet chemistry analysis, after eliminating outlier values which were considered results of experimental error. Outliers were values $> 1.5\%$, in the case of glucose, and $> 2.0\%$ in the case of sucrose as these are the limits of both sugars in almost all potato cultivars (Storey, 2007). Mean and standard deviation values are higher in the 2011 season than 2009 season which is a result of lower temperature (1°C) and the fewer number of samples. Moreover, skewness resulted in both seasons especially in the case of sucrose even though the experiment was designed to minimize such. Maximum values of glucose and sucrose obtained from the 2011 season were higher than values in 2009 which is a direct result of the lower temperature (1°C) added to the 2011 season.

Table 4.1. Statistical summary of reference analysis resulted from wet chemistry for Frito Lay 1879 and Russet Norkotah cultivars.

Statistics	2009 Season				2011 Season			
	FL		RN		FL		RN	
	GL (%)	SU (%)	GL (%)	SU (%)	GL (%)	SU (%)	GL (%)	SU (%)
Minimum	0.0028	9.1e-5	0.0031	0.0045	0.0229	0.0031	0.1719	0.0111
Maximum	0.1514	0.1607	0.3574	0.4205	0.2618	0.2999	1.1663	2.2271
Mean	0.0457	0.0330	0.0591	0.10253	0.1016	0.0729	0.5454	0.2904
Median	0.0391	0.0275	0.0338	0.0836	0.0893	0.0611	0.5528	0.1674
Standard Deviation	0.0281	0.0239	0.0688	0.0806	0.0536	0.0560	0.1895	0.3361
Skewness	1.4003	2.819	2.0412	1.2472	0.9401	1.5135	0.2741	2.2217
Kurtosis	6.1725	17.1841	6.0141	4.6476	3.3262	5.6421	3.0054	9.8636

4.3.2 Mean Reflectance Spectra (MRS)

Fig. 4.8 shows the mean reflectance spectra for the sample set of both cultivars at the wavelength range of 400-1000 nm for the 2009 season and 457-973 nm (216 wavelengths) for the 2011 season. The amplitude and shape of the relative mean reflectance for both cultivars are similar with absorption in the visible range at 570 nm possibly due to the yellow color (Penner, 2003). Another absorption peak in the NIR range, around 876 nm, is possibly occurring due to C-H fundamental bands or their combination (Workman and Weyer, 2008).

Similar trend was noted within the 2011 season except for a considerable change being the absorption peaks were shifted from the 2009 season case, and located at 540 nm, and 920 nm. Also, amplitude values for the mean reflectance in 2011 is less than for the 2009 season which can be explained by variations in performance of differing systems.

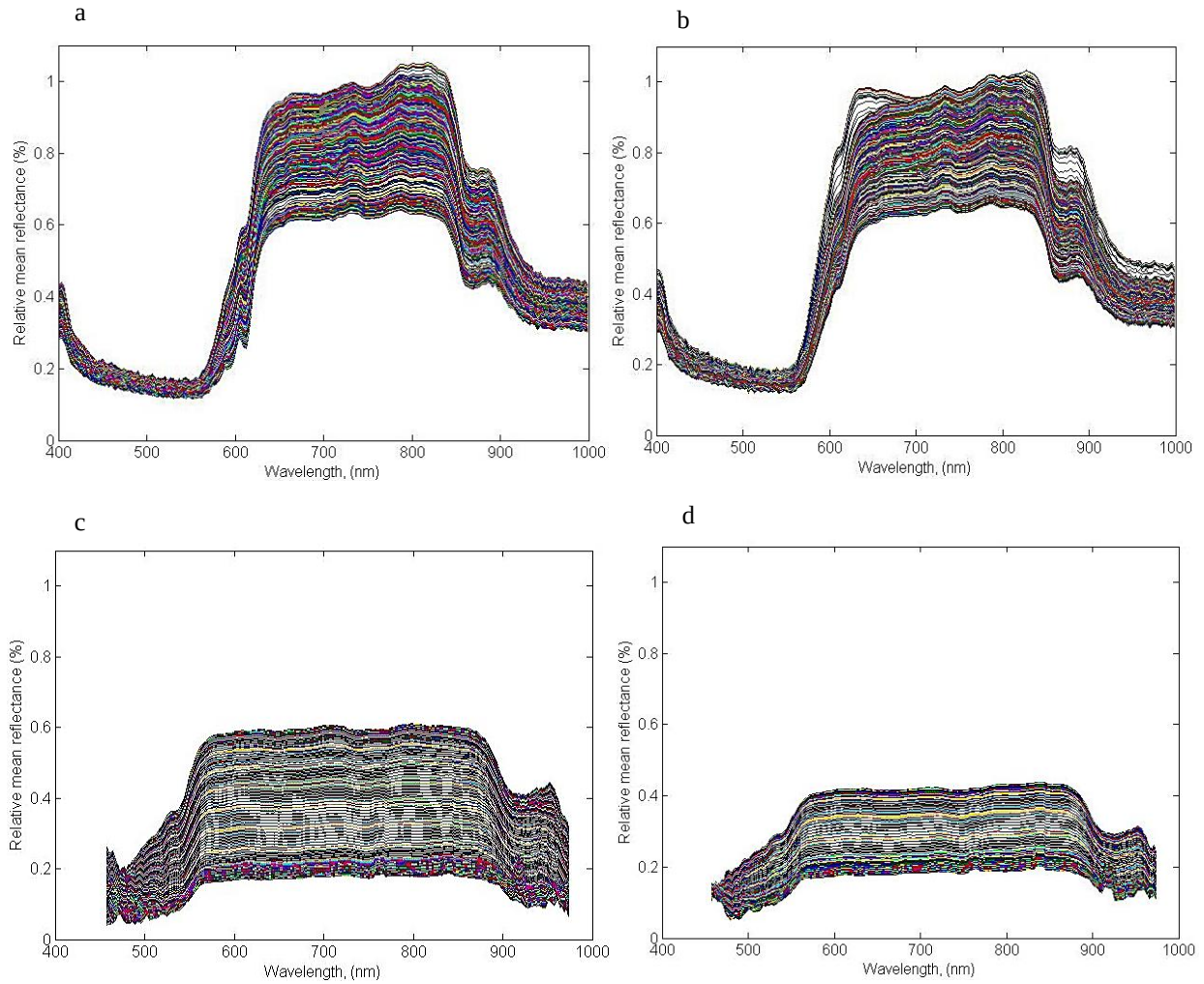


Figure 4.8. Relative mean reflectance for a. Frito Lay 1879, b. Russet Norkotah, for the 2009 season, and relative mean reflectance for c. Frito Lay 1879, and d. Russet Norkotah, for the 2011 season.

4.3.3 Curve Fitting Parameters

Using the exponential model, curve-fitting parameter a_{wi} , which represents the maximum intensity value, was estimated and normalized using Teflon material. In the 2009 season, only the wavelength range 550-1000 nm (192 wavelengths) was considered as shown in Fig. 4.9a-b; and that range was 493-973 nm (201 wavelengths) for the 2011 season as shown in Fig. 4.9c-d, because signals beyond these spectral ranges were too noisy. For the 2009 season, an absorption peak was observed at 876 nm with apparent trend as in the mean reflectance spectra explained

before. However, for the 2011 season, FL and RN showed similar trend for a_{wi} with several samples in both cultivars showing two reflectance peaks at 560, and 900 nm with no obvious absorption peaks.

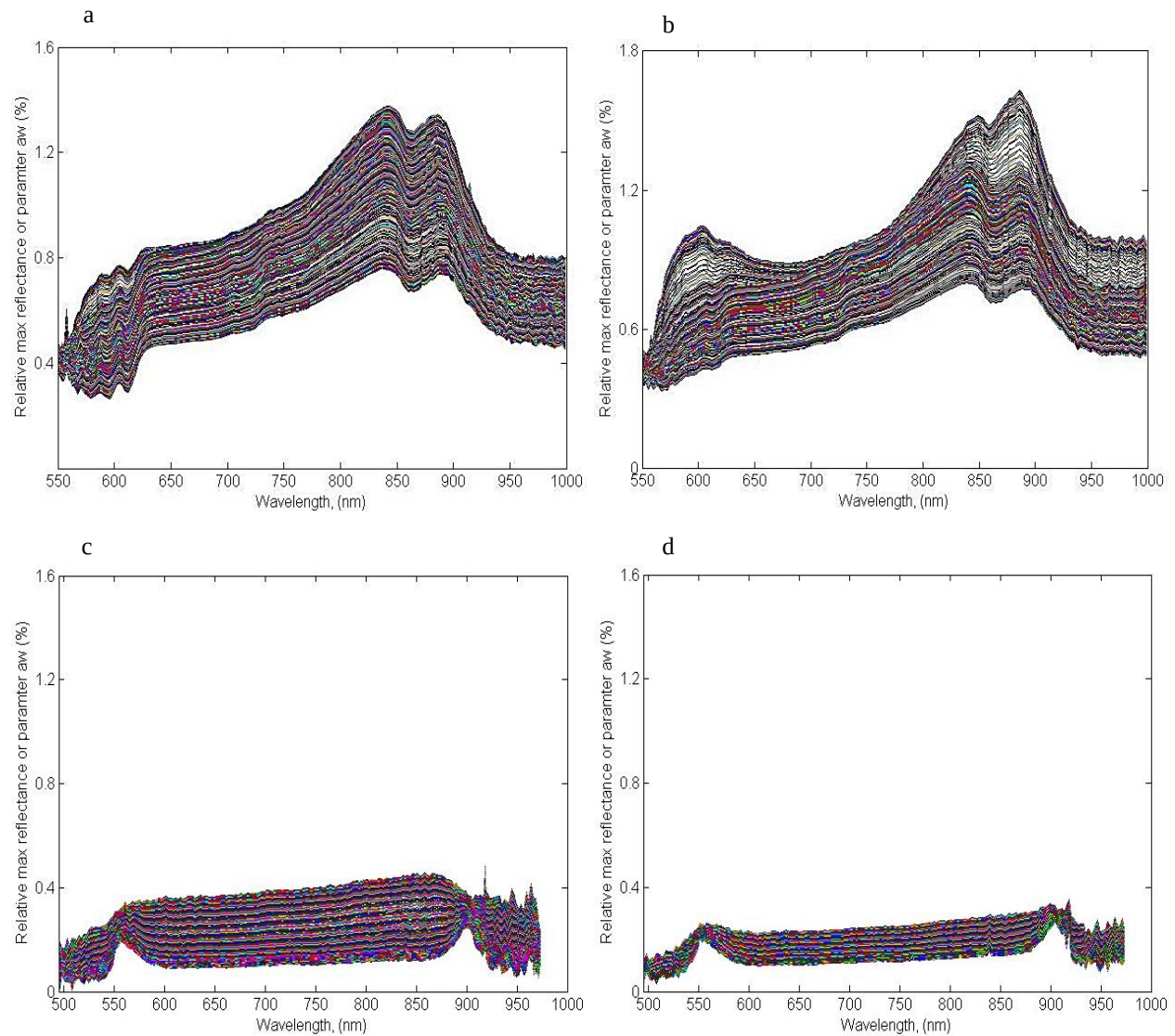


Figure 4.9. Relative parameter a_{wi} for a. Frito Lay 1879, b. Russet Norkotah, for the 2009 season, and relative parameter a_{wi} for c. Frito Lay 1879, and d. Russet Norkotah, for the 2011 season.

In the case of the full width at half maximum (FWHM) of intensity, or b_{wi} , plots are shown at Fig. 4-10a-d for the 2009 season at 586-1000 nm (177 wavelengths) and the 2011

season at 493-973 nm (201 wavelengths). In Fig. 10a-b, there was a peak at 876 nm that showed the maximum value of FWHM for both cultivars for the 2009 season. For the 2011 season, a growing behavior of b_{wi} was observed with absorption peaks at 560 nm and 920 nm with apparent similar trends as in mean reflectance spectra.

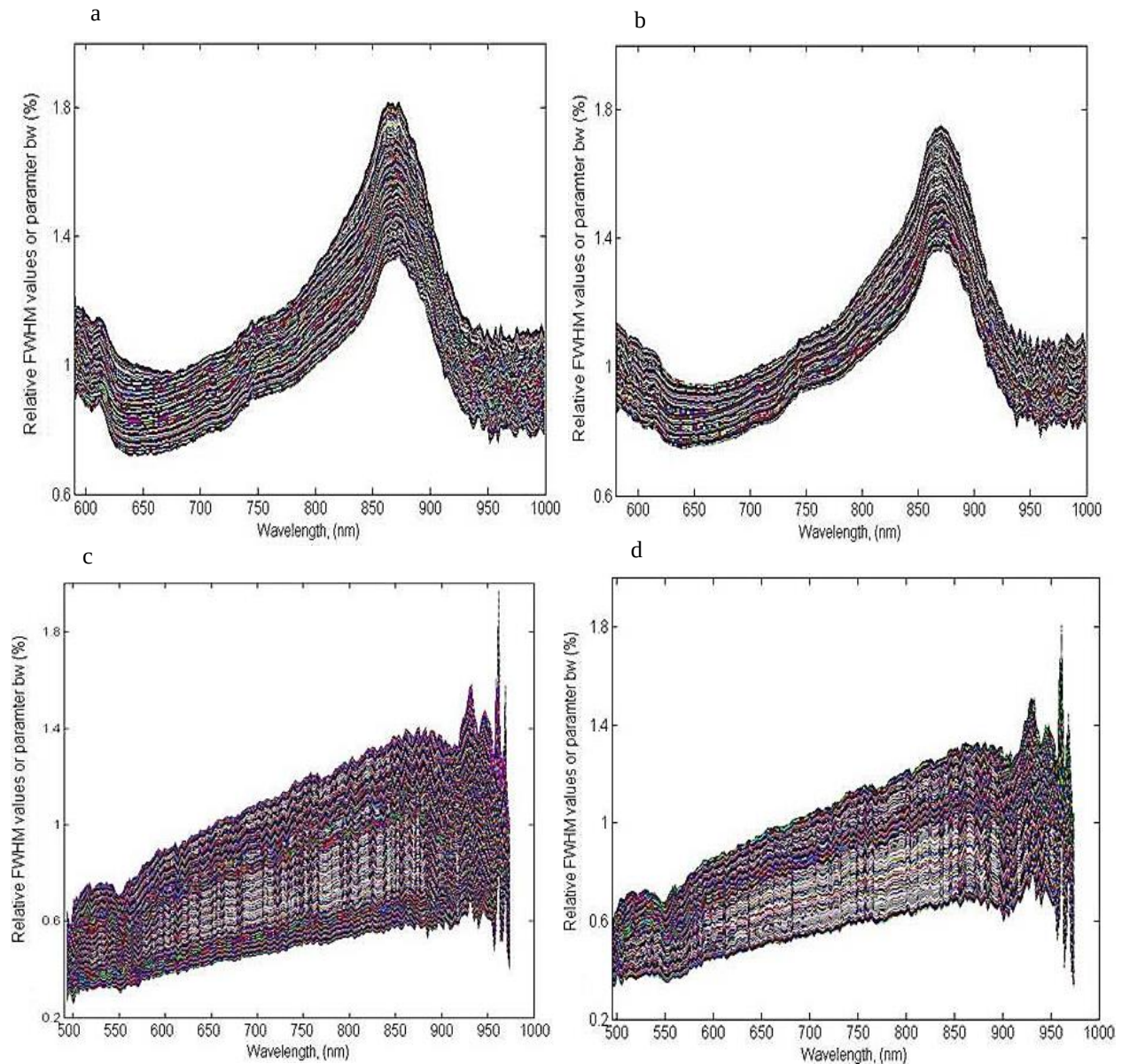


Figure 4.10. Relative parameter b_{wi} for a. Frito Lay 1879, b. Russet Norkotah, for the 2009 season, and relative parameter b_{wi} for c. Frito Lay 1879, and d. Russet Norkotah, for the 2011 season.

4.3.4 Note About Performance of the Hyperspectral System Used in 2011 the Season

After reviewing the difference between the two systems used in the 2009 and 2011 seasons, it was concluded that the results of data obtained in the 2011 season were not accurate and were concerning compared with that in 2009 season as in the latter system, the results are very similar to those obtained from the study in chapter 3. Thus, no further analysis of the previous features was conducted for the 2011 season, and consequently only results of the 2009 are shown.

4.3.5 Partial Least Squares Regression (PLSR) Results

The best results of PLSR for both potato cultivars in the case of mean reflectance spectra (MRS) and combined parameters (concatenated a_{wi} with b_{wi} ; concatenated MRS, a_{wi} , and b_{wi}) for the 2009 season are shown in table 4.2. Results were close between all three data sets (MRS; a_{wi} , and b_{wi} ; MRS, a_{wi} , and b_{wi}). For glucose prediction, RN had stronger correlation than FL with R(RPD) values of 0.96(3.29) for RN and 0.81(1.70) for FL using the MRS, a_{wi} , and b_{wi} combined data set in both cases. For sucrose, however, FL had stronger prediction models than RN with best performance obtained having R(RPD) values as of 0.58(1.23) for FL and 0.30(0.98) for RN and using the a_{wi} and b_{wi} combined data set in both cases. The relationship between measured (actual) and PLSR predicted glucose concentrations, in g/100g of fresh tuber weight, obtained from prediction models for FL and RN is shown in Fig. 4.11.

Table 4.2. PLSR results of predicting glucose and sucrose using VIS/NIR hyperspectral imaging for sliced potato samples in the 2009 season using Frito Lay 1879 and Russet Norkotah cultivars.

Descriptive variable	Cultivar _{Constituent}	Preprocessing ^a	Calibration model				Validation model		
			R _{cal}	RMSEC	RMSEC _{cv} (%)	LVs	R _{pred}	RMSEP (%)	RPD
MRS	FL _{GL}	A ₁₀ ,B ₃ ,C ₀	0.86	0.0158	0.0204	13	0.80	0.0184	1.67
	FL _{SU}	A ₁₀ ,B ₁ ,C ₀	0.71	0.0239	0.0280	10	0.53	0.0282	1.17
	RN _{GL}	A ₀ ,B ₁ ,C ₀	0.97	0.0229	0.0266	10	0.96	0.0289	3.21
	RN _{SU}	A ₀ ,B ₁ ,C ₀	0.41	0.0788	0.0860	5	0.27	0.0694	0.97
a _{wi} , b _{wi}	FL _{GL}	A ₀ ,B ₃ ,C ₀	0.81	0.0182	0.0202	19	0.78	0.0192	1.60
	FL _{SU}	A ₀ ,B ₁ ,C ₀	0.68	0.0248	0.0269	15	0.58	0.0268	1.23
	RN _{GL}	A ₀ ,B ₁ ,C ₀	0.97	0.0235	0.0285	19	0.96	0.0285	3.26
	RN _{SU}	A ₀ ,B ₁ ,C ₀	0.41	0.0790	0.0855	7	0.30	0.0686	0.98
MRS, a _{wi} , b _{wi}	FL _{GL}	A ₀ ,B ₁ ,C ₀	0.85	0.0166	0.0195	20	0.81	0.0181	1.70
	FL _{SU}	A ₀ ,B ₁ ,C ₀	0.70	0.0241	0.0278	12	0.56	0.0274	1.21
	RN _{GL}	A ₀ ,B ₁ ,C ₀	0.97	0.0233	0.0269	11	0.96	0.0282	3.29
	RN _{SU}	A ₆ ,B ₃ ,C ₀	0.39	0.0798	0.0843	5	0.30	0.0684	0.98

^a See table 3.1 footnote.

MRS: mean reflectance spectra, FLGL: Frito Lay 1879, glucose, FLSU: Frito Lay 1879, sucrose, RNGL: Russet Norkotah glucose, RNSU: Russet Norkotah, sucrose.

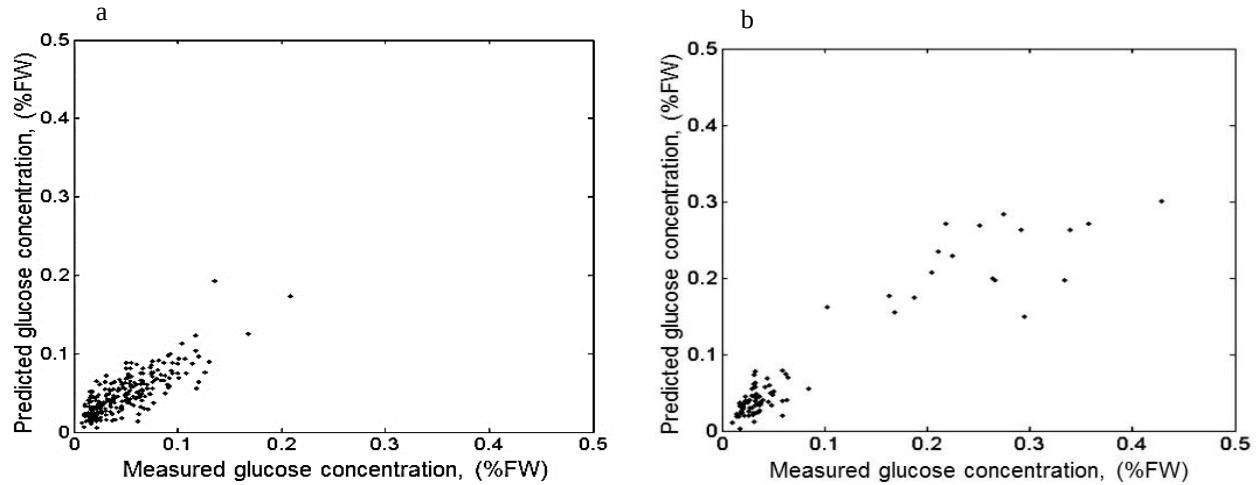


Figure 4.11. Relationship between measured and predicted glucose values for sliced samples using full wavelengths for a) Frito Lay1879 and b) Russet Norkotah cultivars in the 2009 season using PLSR as indicated in table 4.2.

4.3.6 Artificial Neural Network (ANN) Results

The results of artificial neural network used to predict glucose and sucrose sugars for the 2009 season are shown in table 4.3. Compared with PLSR results, lower performance was achieved except for glucose prediction models for RN. FFNN yielded the best prediction models for glucose in the 2009 season with R(RPD) values as high as 0.96(3.05) obtained from

combined a_{wi} and b_{wi} data. Also FL glucose prediction models showed values of R(RPD) as high as 0.74(1.48) obtained from combining MRS, a_{wi} , and b_{wi} . Sucrose prediction models showed slightly less performance compared with PLSR models with values of R(RPD) of 0.51(1.11) using FFNN for FL and 0.18(0.65) for RN resulted from combined a_{wi} and b_{wi} data using RBFNN. Fig. 4.12 shows correlation between measured and ANN-predicted glucose concentrations using the test set of data for both FL and RN.

Table 4.3. Results of prediction models to predict glucose and sucrose for sliced potato samples tested by VIS/NIR hyperspectral imaging and using RBFNN, RBFNNE, and FFNN in the 2009 season.

Descriptive variable	Cultivar ^a Constituent	ANN type, characteristics	Training			Testing		
			R_{train}	SeCV _{train} (%)	RPD	R_{test}	Se _{test} (%)	RPD
MRS	FL _{GL}	FFNN, 100	0.75	0.0212	1.37	0.73	0.0212	1.46
	FL _{SU}	FFNN, 100	0.37	0.0261	1.05	0.30	0.0284	0.99
	RN _{GL}	FFNN, 300	0.96	0.0296	3.38	0.94	0.0348	2.56
	RN _{SU}	RBFNN	0.32	0.1098	0.72	0.18	0.1083	0.65
a_{wi} , b_{wi}	FL _{GL}	FFNN, 250	0.75	0.0212	1.37	0.72	0.0216	1.43
	FL _{SU}	FFNN, 50	0.53	0.0288	1.17	0.51	0.0261	1.11
	RN _{GL}	FFNN, 200	0.97	0.0262	3.83	0.96	0.0291	3.05
	RN _{SU}	RBFNN	0.32	0.1098	0.72	0.18	0.1083	0.65
MRS, a_{wi} , b_{wi}	FL _{GL}	FFNN, 250	0.79	0.0179	1.63	0.74	0.0209	1.48
	FL _{SU}	FFNN, 200	0.39	0.0265	1.04	0.30	0.0277	1.02
	RN _{GL}	FFNN, 200	0.95	0.0318	3.15	0.94	0.0328	2.71
	RN _{SU}	RBFNN	0.30	0.1060	0.77	0.13	0.1057	0.63

^a See table 3.1 footnote.

MRS: mean reflectance spectra, FLGL: Frito Lay 1879, glucose, FLSU: Frito Lay 1879, sucrose, RNGL: Russet Norkotah glucose, RNSU: Russet Norkotah, sucrose.

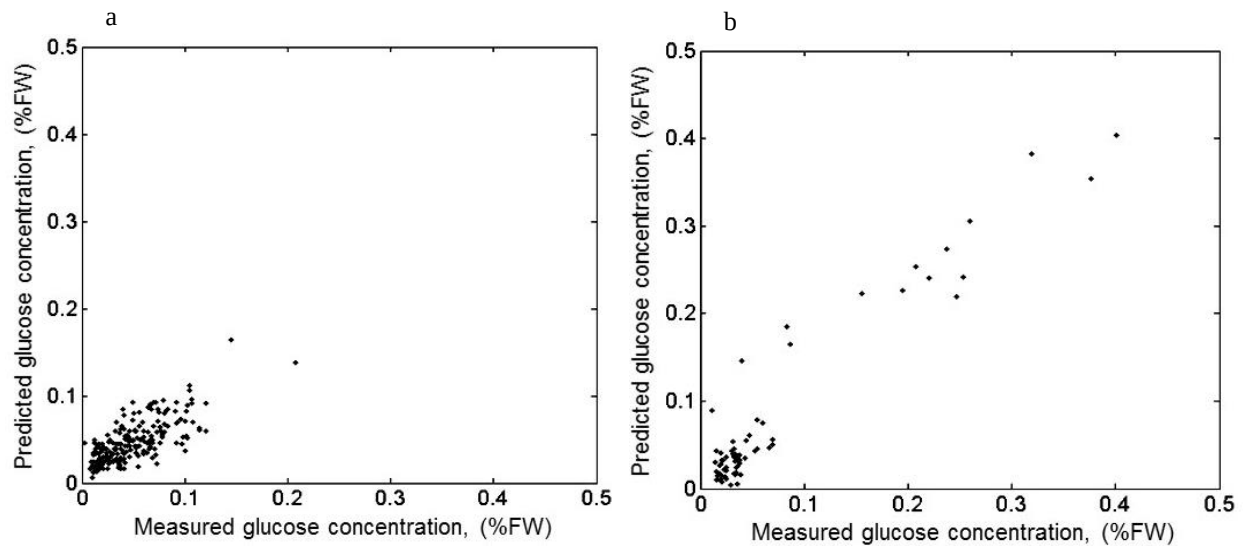


Figure 4.12. Relationship between measured and predicted glucose values for sliced samples using full wavelengths for a) Frito Lay1879 and b) Russet Norkotah in the 2009 season using ANN as indicated in table 4.3.

4.3.7 Variable Selection Results

After observing the closeness between the results obtained by the three data sets used for building prediction models using PLSR and ANN, wavelength selection using IPLS and GA techniques was only based on MRS data. The number of the most effective wavelengths were obtained as shown in table 4.4, both PLSR and ANN were then applied on the selected wavelengths to build prediction models and compared with the full variables models.

Table 4.4. Wavelength selection results using IPLS and GA in the case of glucose and sucrose for potato sliced samples tested VIS/NIR by hyperspectral imaging and in the 2009 season for Frito Lay 1879 and Russet Norkotah.

Selection method	Cultivar _{Constituent}	No. of selected wavelengths in VIS range	No. of selected wavelengths in NIR range
IPLS	FL _{GL}	21	19
	FL _{SU}	10	0
	RN _{GL}	9	3
	RN _{SU}	7	0
GA	FL _{GL}	82	44
	FL _{SU}	78	51
	RN _{GL}	75	39
	RN _{SU}	61	30

In the case of PLSR for 2009 samples as presented in table 4.5, for FL glucose prediction models, R(RPD) were 0.80(1.68) for FL and 0.97(3.66) for RN using IPLS and GA respectively. Sucrose prediction models showed R(RPD) values of 0.54(1.17) and 0.38(1.00) for RN using GA and IPLS respectively. Such results for the 2009 season are similar or slightly better compared with full variables results which indicates the effectiveness of the detected wavelengths clarified in table 4.4. It should be noted that IPLS results in less selected variables than GA which gives it the priority of selection over GA.

Table 4.5. PLSR results for predicting glucose and sucrose using VIS/NIR hyperspectral imaging and selected wavelengths obtained by IPLS and GA for sliced samples in the 2009 season for Frito Lay 1879 and Russet Norkotah cultivars.

Variable selection method	Cultivar _{constituent}	Preprocessing ^a	Calibration model				Validation model		
			R _{cal}	RMSEC	RMSEC _{cv}	LVs	R _{pred}	RMSEP	RPD
IPLS	FL _{GL}	A ₀ ,B ₁ ,C ₀	0.82	0.0176	0.0190	19	0.80	0.0183	1.68
	FL _{SU}	A ₀ ,B ₁ ,C ₀	0.58	0.0233	0.0251	18	0.52	0.0258	1.16
	RN _{GL}	A ₀ ,B ₁ ,C ₀	0.98	0.0221	0.0260	20	0.96	0.0261	3.56
	RN _{SU}	A ₀ ,B ₁ ,C ₀	0.46	0.0769	0.0823	5	0.38	0.0668	1.00
GA	FL _{GL}	A ₀ ,B ₁ ,C ₀	0.82	0.0176	0.0196	14	0.79	0.0190	1.62
	FL _{SU}	A ₀ ,B ₁ ,C ₀	0.72	0.0234	0.0269	14	0.54	0.0281	1.17
	RN _{GL}	A ₀ ,B ₁ ,C ₀	0.98	0.0209	0.0255	14	0.97	0.0254	3.66
	RN _{SU}	A ₀ ,B ₁ ,C ₀	0.41	0.0789	0.0835	5	0.33	0.0676	0.99

^a See table 3.1 footnote.

MRS: mean reflectance spectra, FLGL: Frito Lay 1879, glucose, FLSU: Frito Lay 1879, sucrose, RNGL: Russet Norkotah glucose, RNSU: Russet Norkotah, sucrose.

Results of artificial neural network prediction models after applying variable selection using IPLS and GA for the 2009 season are shown in table 4.6. For FL glucose prediction, values of R(RPD) for RN were as high as 0.96(3.04) and for FL the values were 0.73(1.46) obtained from FFNN using GA in both cases. In the case of sucrose prediction for FL, there was no improvement of correlation compared with PLSR or ANN for full models. FFNN was proven to produce such performance in both cases. In general, the number of selected variables using GA is more than IPLS with close results between the prediction models performance using ANN. Thus, based on computation times, IPLS showed more efficiency than GA in the prediction of glucose and sucrose.

Table 4.6. Artificial neural network results for predicting glucose and sucrose using VIS/NIR hyperspectral imaging and selected wavelengths obtained by IPLS and GA for sliced samples in the 2009 season for Frito Lay 1879 and Russet Norkotah cultivars.

Descriptive variable	Cultivar _{constituent}	ANN type, characteristics	Training			Testing		
			R _{train}	SeCV _{train} (%)	RPD	R _{test}	SeCV _{test} (%)	RPD
IPLS	FL _{GL}	FFNN, 100	0.75	0.0190	1.53	0.70	0.0221	1.40
	FL _{SU}	FFNN, 50	0.26	0.0295	1.14	0.23	0.0290	1.00
	RN _{GL}	FFNN, 50	0.96	0.0284	3.53	0.95	0.0325	2.73
	RN _{SU}	FFNN, 50	0.24	0.0689	1.15	0.13	0.0701	1.00
GA	FL _{GL}	FFNN, 50	0.79	0.0170	1.71	0.73	0.0196	1.46
	FL _{SU}	FFNN, 50	0.26	0.0296	1.14	0.25	0.0294	0.98
	RN _{GL}	FFNN, 150	0.97	0.0250	4.01	0.96	0.0293	3.04
	RN _{SU}	RBFNN	0.21	0.1192	0.67	0.20	0.1041	0.67

MRS: mean reflectance spectra, FLGL: Frito Lay 1879, glucose, FLSU: Frito Lay 1879, sucrose, RNGL: Russet Norkotah glucose, RNSU: Russet Norkotah, sucrose.4

4.3.8 Potatoes Classification Based on Sugar Levels

The numbers of samples in class 1 (less than threshold) and class 2 (above threshold) for glucose and sucrose in the case of both cultivars are shown in table 4.7.

Table 4.7. Numbers of samples in each class based on glucose and sucrose levels for the 2009 season in the case of Frito Lay1879 and Russet Norkotah cultivars.

Cultivar _{constituent}	Class 1 (less than threshold)	Class 2 (above threshold)
FL _{GL}	453	618
FL _{SU}	393	980
RN _{GL}	188	169
RN _{SU}	198	160

Results of sugar classification of potato sliced samples using Knn and PLSDA for FL and RN are shown in table 4.8. Classification error for the training group based on cross validation (training error) and testing error showed that PLSDA resulted in better performance than Knn with the possible reason being the suitability of PLSDA to cope with collinearity. Testing error for glucose classification for FL and RN were 19% and 18% respectively, whereas for sucrose the values were 34% and 38%. Classification results somewhat match with prediction performance as shown in previous sections in which glucose prediction models resulted in better results than sucrose for both cultivars.

Table 4.8. Classification results of sliced samples based on glucose and sucrose levels for the 2009 season using VIS/NIR hyperspectral imaging for Frito Lay1879 and Russet Norkotah cultivars.

Cultivar _{Constituent}	Preprocessing for Knn ; PLSDA ^a	Training error (%)		Testing error (%)	
		Knn	PLSDA	Knn	PLSDA
FL _{GL}	A ₆ ,C ₀ ; A ₄ ,C ₀	19	16	22	19
FL _{SU}	A ₀ ,C ₀ ; A ₅ ,C ₀	39	32	42	34
RN _{GL}	A ₇ ,C ₀ ; A ₉ ,C ₀	20	16	22	18
RN _{SU}	A ₆ ,C ₀ ; A ₅ ,C ₀	41	36	44	38

^a See table 3.1 footnote.

4.4 Conclusions

Partial least squares regression (PLSR) and artificial neural network (ANN) were used to obtain prediction models for glucose and sucrose sugars in 12.7 mm sliced samples, obtained from Frito Lay1879 (FL) and Russet Norkotah (RN) cultivars, using a hyperspectral imaging system in the reflectance mode in the wavelength range of 400-1000 nm. Prediction models based on mean reflectance spectra (MRS) were shown to be more efficient than models based on spectral curve fitting parameters due to similar performance, and fewer variables contained in MRS. PLSR showed similar performance to ANN for both cultivars with R values being as high as 0.81 and 0.97 for FL and RN in the case of glucose. However, weaker performance was achieved for sucrose, compared to glucose, with R values of 0.58 and 0.27. In general, FL as a chipping cultivar with lower glucose than in RN, yielded weaker prediction models for glucose than RN. Prediction models built using selected wavelengths, by interval partial least squares (IPLS), showed similar performance as the full wavelengths' models for both cultivars for glucose with a slight improvement for sucrose prediction with R values of 0.60, and 0.38 for FL and RN. The selected wavelengths results, which are unique in the study of predicting sugars content of potatoes, demonstrate the possibility of reducing data dimensionality and potentially enhancing prediction results. With broader selection of window size, cross validation, mutation rate, cross over breeding, replicate runs, or step size so that more effective wavelengths are selected, the potential exists for improved results. Moreover, prediction models of sucrose did not result in reliable performance and they are not suitable for industrial applications. Consequently, such models need improvement, which can be achieved by increasing the number of samples and/or using several storage temperatures to obtain broader sugar distribution.

Being a novel application of hyperspectral imaging to build prediction and classification models based on sugars in potatoes, this study in general presented a promising application for constituent monitoring of potatoes that are destined to products sensitive to excessive sugar content (chipping and French fries). With further study of extending this approach to intact whole tubers and with the improvement of hardware components in the hyperspectral system, the on-line sorting for potato tubers is a realistic target. Moreover, it is worth stating that, in the meantime and with the available components in the market, it is possible to benefit from the selected wavelengths for building a multispectral system to overcome the problem of relatively extensive time required for image acquisition related to hyperspectral imaging.

CHAPTER 5 UTILIZATION OF VISIBLE/NEAR-INFRARED SPECTROSCOPIC AND WAVELENGTH SELECTION METHODS IN SUGAR PREDICTION AND POTATOES CLASSIFICATION

(Expanded from Rady, A.M., Guyer, D.E. 2014. Utilization of visible/near-infrared spectroscopic and wavelength selection methods in sugar prediction and potatoes classification. *Journal of Food Measurement and Characterization*, in press)

5.1 Introduction

Near-infrared (NIR) is becoming a promising technology that could be extensively used in quality control and monitoring for chemical, petrochemical, pharmaceutical, agricultural, and food industries. As rapid, and/or noninvasive methods, NIR techniques are suitable for on-line applications which are less time consuming, more robust, more reproducible, and more cost effective than human labor or other laboratory methods used in quality assurance. Fruits and vegetables, as high moisture products and having a relatively big size, were not traditionally suitable for NIR applications. However, with development of high performance hardware, intact fruits and vegetable quality measurements using NIR have become feasible using interactance and transmission modes (Kawano, 2002). NIR interactance mode was developed in a USDA laboratory at Beltsville by Conway et al., 1984 to measure human body fat. Later, the practice of NIR interactance in the field of agriculture became more intensive. Sugar accumulation in potato tubers showed that sugar content in potatoes is influenced by storage conditions (temperature, period), and reconditioning more than pre-harvest practices (soil composition, fertilization, environment, irrigation) (Burton et al., 1992).

This study is initial work toward developing a rapid hand-held device that can be used to assess some constituents in potato tubers which will potentially benefit people working in the

potato industry starting from grower and ending with customer. In the current chapter, three data sets were obtained from the VIS/NIR interactance system; full wavelengths, sampled wavelengths, and selected wavelengths using IPLS and GA. The analysis of such data sets included building prediction models for glucose and sucrose, and classification of sliced samples and whole tubers using various techniques.

The short and long term objectives of this research are:

1. Determine the potential of VIS/NIR interactance spectroscopy for quantifying the level of sucrose and glucose levels in potato tubers.
2. Development of a reliable prediction models that may be used to detect the amount of sucrose and glucose in potato tubers at levels which are important for potato growers and processors.
3. Detect the most effective wavelengths related to glucose and sucrose absorption.
4. Study the potential of classifying potatoes based on sugar levels associated with the frying industry using several classifiers as well as classifier fusion.

5.2 Materials and Methods

5.2.1 Raw Material and Experimental Design

The samples used to conduct experiments in this study were the same as those used in section 4.2.1 as well as the experimental design for both the 2009 and 2011 seasons.

5.2.2 Constituent Measurement

5.2.2.1 Potato sample preparation

In addition to the sliced samples prepared as noted in section 4.2.2.1, whole tubers were also used in electronic measurements. The scan position for the whole tuber was chosen such that the incident light penetrates the area above the tissue extracted for juicing, and located in the middle of both axes. Regarding the spatial variation of a potato surface, the electronic measurements were made such that the area receiving the incident light, for sliced samples and whole tubers, is as uniform as possible for all samples. Consequently, the yielded variation will be due to the differences between samples in light absorption under the surface and not due to the spatial variation of the tuber surface.

5.2.2.2 Wet chemistry basis measurements

The procedure used in juice extraction from sliced samples was the same as that used in section 4.2.2.2.1. Also, wet chemistry steps conducted to evaluate glucose and sucrose concentrations were the same as mentioned in section 4.2.2.2 and using equations 3.2 and 3.3 respectively.

5.2.3 VIS/NIR Interactance System

The system used in this study had the same components and configurations as that used in section 3.2.2.4. A standard Teflon[®] as a reference material and then equation 3.1 was applied to obtain the relative absorption.

5.2.4 Data Analysis Discussion and Approach

5.2.4.1 Data handling

Various scenarios of handling and statistically analyzing the relative absorption data extracted from the interactance experiments were conducted. Fig. 5.1 shows the sequence of data handling and methods used to build prediction models. First, the signals resulting from measurements were reduced from 3648 to 2701 wavelengths (from 446-1125 nm to 501-1004 nm) based on visual evaluations. Next, two modes of data were tested, the data containing full 2701 wavelengths, and sampling at every 7 wavelengths resulting in 386-variable matrices. Finally, data from the 2009 and 2011 seasons were combined for both the full and the sampled variables. Both PLSR and ANN were applied to each data set to obtain prediction models for both glucose and sucrose.

5.2.4.2 Partial least squares regression (PLSR)

A complete description of PLSR used in this research along with pretreatment for either spectra or reference values is listed in section 3.2.4.

5.2.4.3 Artificial neural network (ANN)

The same artificial neural network types, and configuration applied in this study were the same as that used in section 4.2.4.3.

5.2.4.4 Wavelength selection

A complete description of interval partial least squares (IPLS) and genetic algorithm (GA) used in this research was listed in section 4.2.4.5. Variable selection techniques for interactance data were only applied on the sampled data (386) as the number of full variables

(wavelengths) is 2701 and it's not possible to conduct variable selection on this case using either IPLS or GA.

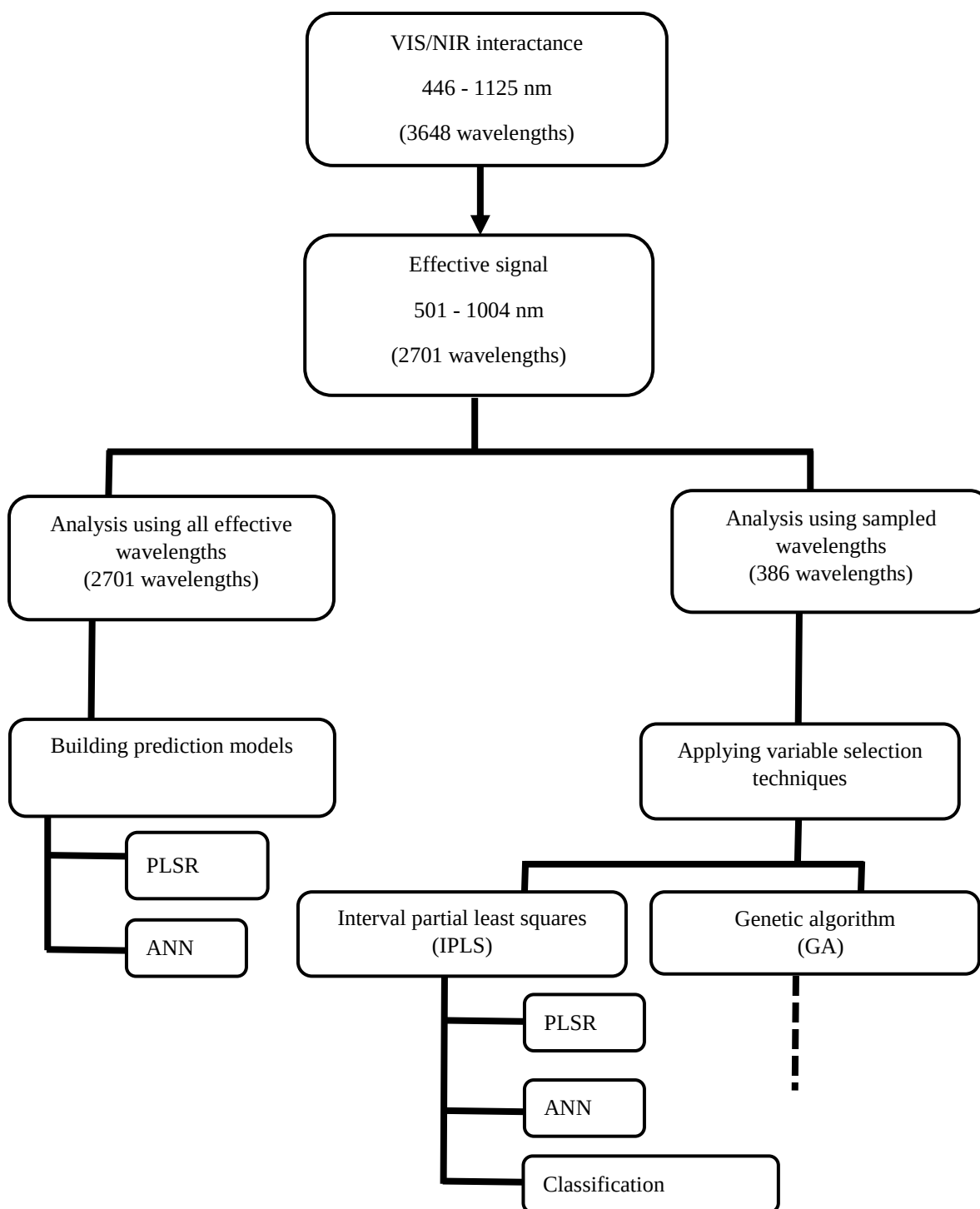


Figure 5.1. A schematic diagram of data handling and analysis for data obtained using VIS/NIR interactance spectroscopy to obtain prediction models of glucose and sucrose and for classification of Frito Lay1879 and Russet Norkotah based on sugar levels.

5.2.4.5 Classification of potatoes based on sugar levels

Classification of whole tubers and sliced samples based on glucose and sucrose levels was conducted as described in section 4.2.4.6. Moreover, several techniques were applied to enhance classification accuracy. In addition to Knn and PLSDA, linear discriminant analysis (LDA), and artificial neural network (ANN) were also used for the classification. In discriminant analysis, each sample is assigned to a class. For LDA, the decision boundary is a hyper plane that separates the two classes (Bishop, 2007; Duda, et. al., 2001). In the current study, Euclidean, as well as Mahalanobis, distances were applied for assigning each sample to the appropriate class. Only principal components (20 components that were responsible for >99% of the total variance) were used for LDA as they overcome the problem of colinearity associated with spectroscopic measurements.

ANN classification was based on FFNN that consisted of an input layer containing the pretreated spectra data, a hidden layer with 50 neurons, and an output layer that contained the assigned class. Transfer functions were chosen as log-sigmoid, and scaled conjugate gradient back propagation for hidden and output layers respectively. Samples in both seasons were divided into two classes based on the cut-off glucose values in the 2009 season of 0.035% for both FL and RN, whereas the values for sucrose were 0.03% and 0.10%. In the 2011 season, and based on sugar distribution, the threshold values for glucose were 0.09% and 0.5% for FL and RN, while the values for sucrose were 0.08% and 0.15%. Cut-off levels were adopted from recommended thresholds listed by Stark and Love (2003), for both sugars except for the glucose level for RN which was chosen to create two balanced classes. Classification of sugars was conducted using the Matlab[®] statistical toolbox for LDA, and ANN.

Classifier fusion was also conducted to increase the overall classification accuracy. Weighted majority voting was used for setting each sample in the correct class. In majority voting, and based on results obtained from individual classifiers, PLSDA was given the highest weight of 0.40, and weights of 0.20, 0.10, and 0.15 were given to LDA, and Knn, and ANN respectively. Each sample was assigned to the class having the higher total voting resulted from all classifiers.

5.3 Results and Discussions

5.3.1 Constituents Distribution

The basic statistics for both glucose and sucrose over the 2009 and 2011 data were shown in table 4.1. Moreover, sample distributions of glucose and sucrose from wet chemistry for FL and RN in the 2009 and 2011 seasons are shown in Fig. 5.2, with a broader range of both constituents in the 2011 season due to more storage temperatures utilized.

5.3.2 Spectra for Sliced Samples and Whole Tubers

The signals extracted from the VIS/NIR interactance measurement experiments for both cultivars for sliced and whole samples in the range 501-1004 nm, extracted from the original wavelengths signal (446-1125 nm), are shown in Figs 5.3, and 5.4 for the 2009 and 2011 seasons. In general, the signals from whole samples appear less scattered than with sliced samples with peak values of relative interactance being one third of the peak value for the sliced samples indicating the effect of sample preparation (i.e. skin effect) on interactance. In the 2011 season, the same trend was obtained for both cultivars in the case of sliced samples or whole tubers with more condensed signals for whole tubers.

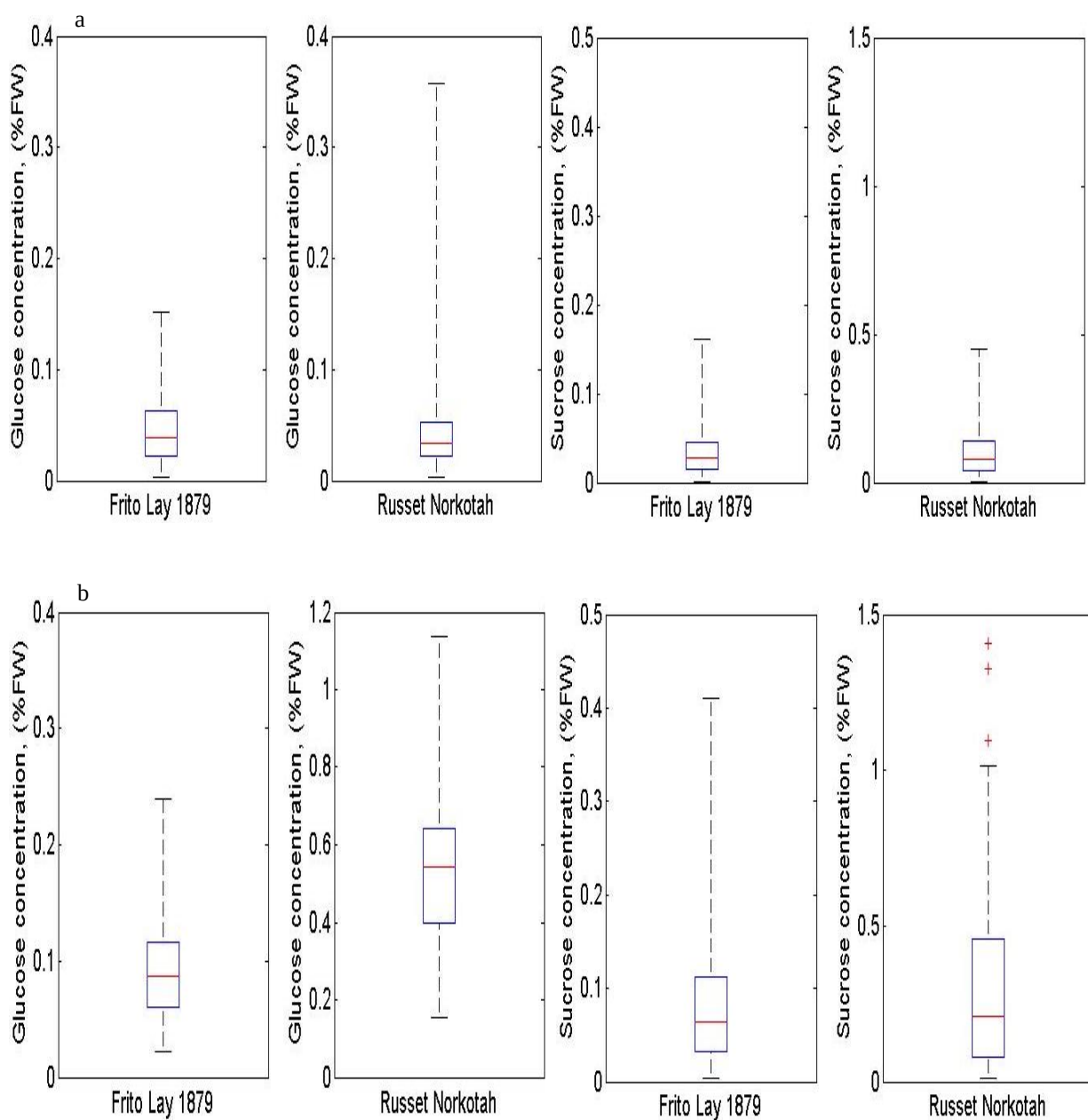


Figure 5.2. Distribution of glucose and sucrose (%FW) for Frito Lay 1879 and Russet Norkotah from wet chemistry in a) 2009, and b) 2011 seasons. Note: scale change on RN glucose for display purpose.

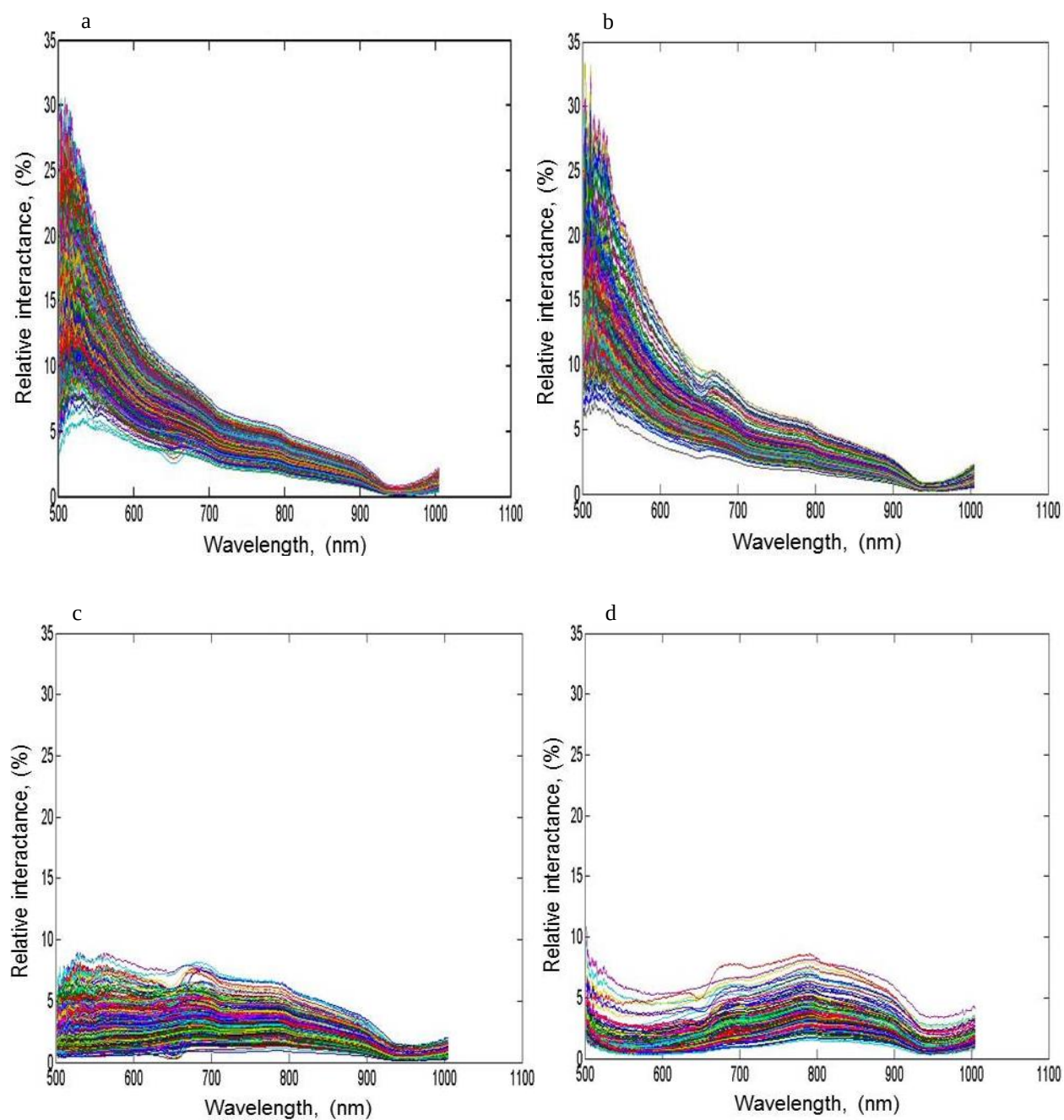


Figure 5.3. Relative intertance of the 2009 season data for sliced samples a. Frito Lay 1879, b. Russet Norkotah, and relative intertance for whole tubers for c. Frito Lay 1879, and d. Russet Norkotah.

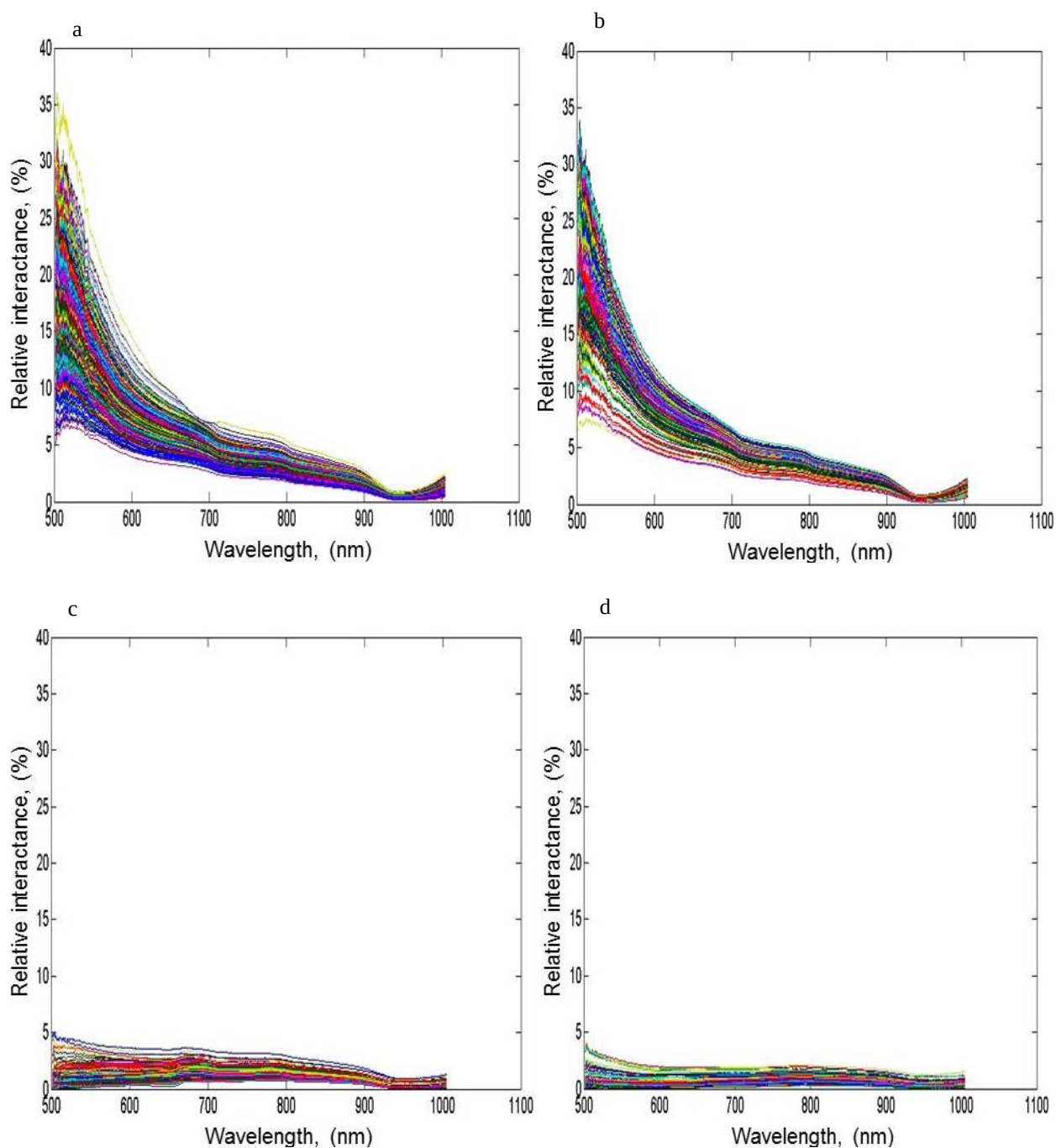


Figure 5.4. Relative intertactance of the 2011 season data for sliced samples a. Frito Lay 1879, b. Russet Norkotah, and relative intertactance for whole tubers for c. Frito Lay 1879, and d. Russet Norkotah.

To obtain more information about the trend of sliced samples and whole tubers under the applied intertactance experiments, the mean spectra of $\log(1/\text{intertactance})$ was calculated and plotted in Fig. 5.5a-b for the 2009 season, where A is the relative intertactance. In the case of

sliced samples, FL and RN showed similar trend in both visible and near-infrared regions with an absorption peak at round 960 nm that is related to OH-water overtone (Chen et al., 2004; Helgerud et al., 2012). For whole tubers, while both cultivars showed similar trends to sliced samples in the NIR region, different behavior in the visible region was observed. Such variation is due to color differences and non-uniformity of the skin surface between FL and RN. An absorption band was noted for RN at 550-600 nm which possibly refers to the absorption of green (490-580 nm), and yellow (580-600 nm) colors, and a slight peak around 650 nm which possibly refers to the absorption of orange (600-650 nm) color. While FL showed two small absorption peaks around 550 nm (green color), and 650 nm (orange color) which also refers to the absorption of the green and orange colors respectively (Giambattista et al., 2007). In the 2011 season (figures are not shown), the same trend was obtained for both cultivars in the case of both sample types.

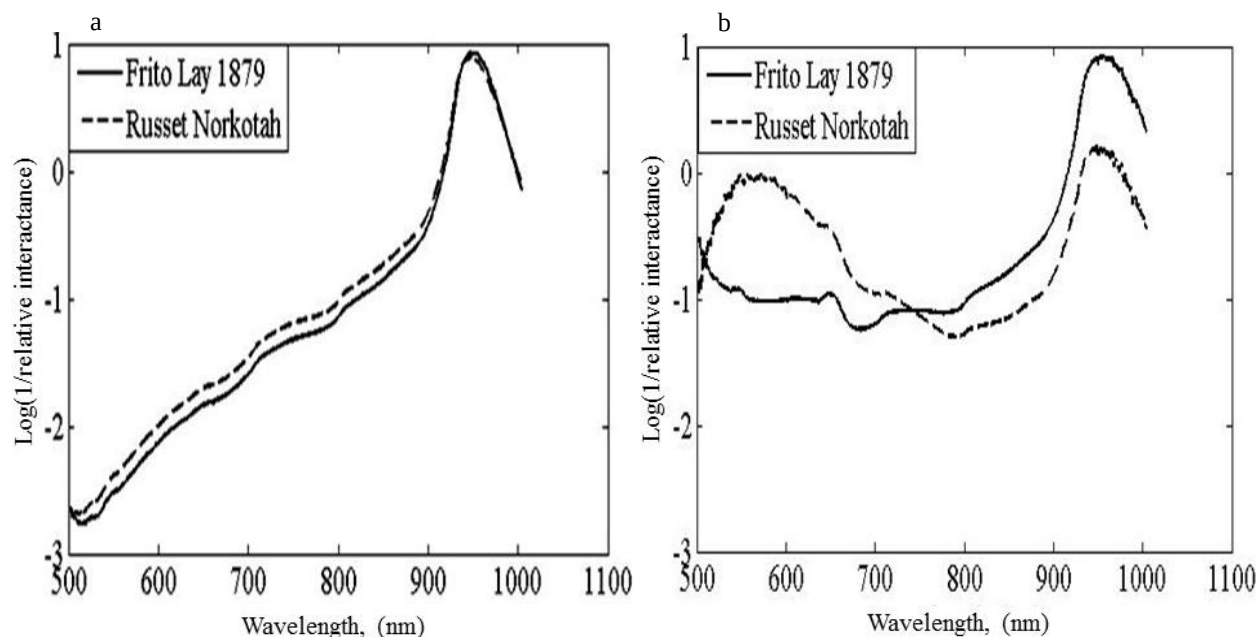


Figure 5.5. Mean of log (1/relative interactance) of the 2009 season data for Frito Lay 1879 and Russet Norkotah for: a. Sliced samples, b. Whole tubers.

5.3.3 Partial Least Squares Regression (PLSR) Results

5.3.3.1 Full and sampled wavelengths models

Results for calibration and prediction models of glucose and sucrose for both Frito Lay1879 (FL) and Russet Norkotah (RN) cultivars in the case of the 2009 and 2011 seasons are shown in table 5.1. In the 2009 season, and based on full wavelengths models, for glucose prediction models, RN yielded strong correlation with R(RPD) values of 0.94(2.85) for sliced samples and 0.97(4.16) for whole tubers. Compared to glucose models, weaker correlation was obtained for sucrose with R(RPD) values of 0.53(1.18) and 0.53(1.16) for sliced samples and whole tubers respectively. In the case of FL, whole tubers yielded glucose prediction models with R(RPD) values of 0.79(1.62) and those values were slightly better than sliced samples models of 0.76(1.53). However, sucrose prediction models had values of R(RPD) of 0.30(1.04),

and 0.33(1.05) of sliced samples and whole tubers respectively. For the prediction models obtained from sampled wavelengths, glucose prediction models of Frito Lay 1879 in the case of whole tubers showed higher correlation than full wavelengths models with R(RPD) values of 0.85(1.92). Other constituents showed similar performance to the full wavelengths models for both cultivars in the case of both glucose and sucrose which generally clarifies the advantage of reducing data dimension using sampling.

In the 2011 season, general lower correlation was achieved of both sugars in the case of both cultivars than in the 2009 season. Values of R(RPD) for FL in the case of glucose were 0.59(1.17) and those values for RN were 0.53(1.15). For sucrose, R(RPD) values for FL were 0.56(1.16) and 0.33(1.02) for RN. For whole tubers, glucose prediction resulted in R(RPD) values of 0.36(1.08) for FL and 0.62(0.70) for RN, and these values were 0.21(0.98) for FL and 0.45(1.12) for RN in the case of sucrose prediction models. The reduced (sampled) data yielded weaker performance than full wavelengths except for sucrose prediction models for RN with R(RPD) values of 0.69 (1.41) for sliced which is even better than 2009 results indicated before.

Table 5.1. PLSR results for predicting glucose and sucrose for sliced samples and whole tubers using VIS/NIR interactance and using full (2701) and sampled wavelengths in the 2009 and 2011 seasons for Frito Lay 1879 and Russet Norkotah cultivars.

Season	Wavelengths	Sample type	Cultivar _{constituent}	Preprocessing ^a	Calibration model				Validation model		
					R _{cal}	RMSEC	RMSEC _{cv} (%)	LVs	R _{pred}	RMSEP (%)	RPD
2009	Full	Slice	FL _{GL}	A ₇ B ₁ C ₀	0.77	0.0178	0.0187	14	0.76	0.0181	1.53
			FL _{SU}	A ₆ B ₁ C ₁	0.33	0.7988	0.8241	13	0.30	0.8167	1.04
			RN _{GL}	A ₉ B ₁ C ₂	0.95	0.0341	0.0364	8	0.94	0.0387	2.85
			RN _{SU}	A ₉ B ₁ C ₀	0.64	0.0613	0.0781	18	0.53	0.0682	1.18
		Whole	FL _{GL}	A ₀ B ₁ C ₀	0.85	0.0149	0.0195	20	0.79	0.0172	1.62
			FL _{SU}	A ₅ B ₁ C ₁	0.37	0.7873	0.8239	15	0.33	0.8082	1.05
			RN _{GL}	A ₇ B ₃ C ₀	0.99	0.0093	0.0263	2	0.97	0.0179	4.16
			RN _{SU}	A ₇ B ₃ C ₀	0.72	0.0555	0.0753	6	0.53	0.0698	1.16
	Sampled	Slice	FL _{GL}	A ₀ B ₁ C ₀	0.78	0.0171	0.0180	20	0.76	0.0178	1.53
			FL _{SU}	A ₀ B ₁ C ₁	0.37	0.7883	0.8252	17	0.29	0.8197	1.04
			RN _{GL}	A ₇ B ₁ C ₂	0.95	0.0335	0.0371	18	0.93	0.0421	2.61
			RN _{SU}	A ₄ B ₃ C ₀	0.67	0.0591	0.0788	20	0.52	0.0692	1.17
		Whole	FL _{GL}	A ₇ B ₃ C ₀	0.89	0.0126	0.0158	4	0.85	0.0142	1.92
			FL _{SU}	A ₄ B ₃ C ₁	0.53	0.7176	0.8115	17	0.35	0.8111	1.05
			RN _{GL}	A ₀ B ₁ C ₀	0.97	0.0153	0.0194	20	0.95	0.0241	3.11
			RN _{SU}	A ₇ B ₃ C ₁	0.74	0.6256	0.9006	8	0.45	0.8745	1.06
2011	Full	Slice	FL _{GL}	A ₇ B ₁ C ₂	0.79	0.0517	0.0746	11	0.59	0.0710	1.17
			FL _{SU}	A ₉ B ₁ C ₀	0.73	0.0476	0.0660	10	0.56	0.0533	1.16
			RN _{GL}	A ₀ B ₃ C ₂	0.87	0.0489	0.0867	7	0.53	0.0914	1.15
			RN _{SU}	A ₀ B ₁ C ₁	0.54	0.9976	1.1514	5	0.33	1.0488	1.02
		Whole	FL _{GL}	A ₀ B ₁ C ₁	0.41	0.4985	0.5280	5	0.36	0.4789	1.08
			FL _{SU}	A ₁₀ B ₁ C ₁	0.42	0.6954	0.7714	6	0.21	0.6813	0.98
			RN _{GL}	A ₆ B ₃ C ₁	0.98	0.0628	0.2875	8	0.62	0.5865	0.70
			RN _{SU}	A ₀ B ₁ C ₁	0.46	0.7909	0.9258	4	0.45	0.8862	1.12
	Sampled	Slice	FL _{GL}	A ₅ B ₁ C ₀	0.48	0.5025	0.5655	10	0.47	0.4939	1.14
			FL _{SU}	A ₆ B ₁ C ₂	0.39	0.1050	0.1145	8	0.26	0.1108	1.03
			RN _{GL}	A ₄ B ₁ C ₁	0.43	0.3784	0.4041	3	0.20	0.4567	1.01
			RN _{SU}	A ₀ B ₁ C ₁	0.36	1.1128	1.1837	3	0.31	1.0131	1.05
		Whole	FL _{GL}	A ₁ B ₁ C ₁	0.48	0.4790	0.5309	5	0.37	0.4849	1.07
			FL _{SU}	A ₁ B ₃ C ₁	0.36	0.6906	0.7818	1	0.28	0.6131	1.04
			RN _{GL}	A ₀ B ₃ C ₀	0.66	0.1362	0.2345	1	0.39	0.1803	1.02
			RN _{SU}	A ₇ B ₁ C ₂	0.70	0.1219	0.1518	5	0.69	0.1406	1.41

^a See table 3.1 footnote.

5.3.3.2 Selected variables-PLSR models

Prediction models for glucose and sucrose in the case of Frito Lay 1879 and RN using PLSR and based on selected variables from both IPLS and GA are shown in table 5.2 for the 2009 and 2011 seasons with the number of selected wavelengths in table 5.3. Comparing between PLSR results for selected-variables models with those obtained from full or sampled wavelengths and for 2009 showed that selected wavelengths-prediction models yielded the same correlation performance, or slightly better, as full wavelengths as well as sampled wavelengths models. Slightly better performance was obtained for FL glucose prediction models for sliced samples and whole tubers with R(RPD) values of 0.79(1.61) and 0.81(1.72) using IPLS. Moreover, an improvement in sucrose prediction models for RN in the case of sliced samples and whole tubers with R(RPD) values of 0.55(1.18) and 0.64(1.30) were obtained from GA and IPLS respectively.

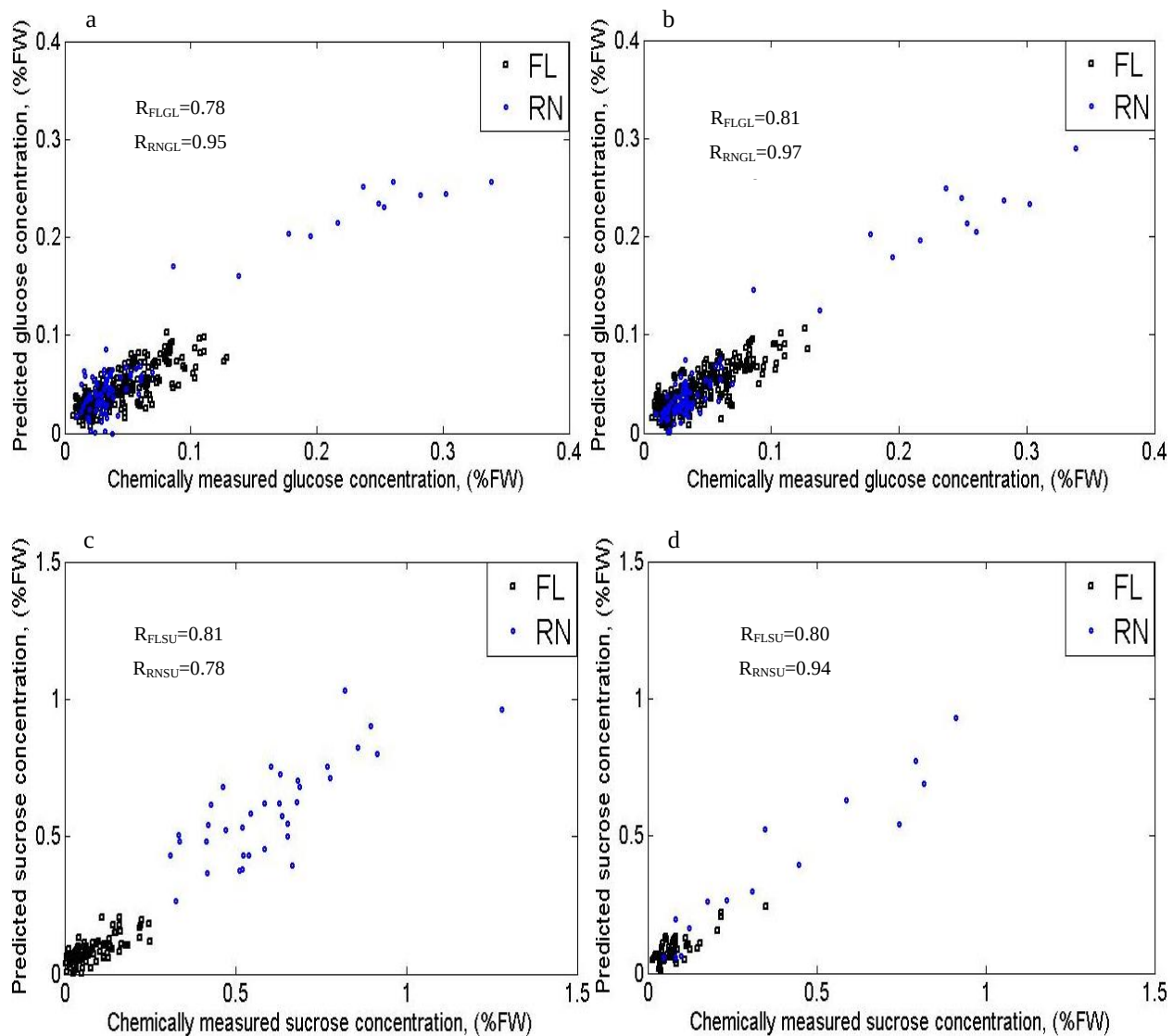
For the 2011 season, significant improvements were obtained compared with full or sampled wavelengths models. For sliced samples, glucose prediction models for FL and RN showed R(RPD) values as high as 0.74(1.49) and 0.88(2.12), obtained from IPLS and GA respectively. Sucrose prediction models also were improved for both cultivars and yielded prediction models with R(RPD) values of 0.81(1.70) for FL using GA and 0.71(1.32) for RN using IPLS. Whole tubers also showed considerable improvements with R(RPD) values for glucose models as high as 0.71(1.32) 0.91(2.08) for FL and RN respectively and using IPLS as a wavelength selection technique in both cases. In the case of sucrose, R(RPD) values were 0.80(1.64) and 0.94(2.82) for FL and RN respectively using IPLS. In general, IPLS yielded better PLSR prediction models, using different window sizes, than GA, that has window size of 1, for both cultivars in 2009 and 2011 data sets

with some exceptions shown in table 5.2. The best PLSR prediction models representing the relationship between measured and predicted values of glucose and sucrose for the 2009 season for are shown in Fig. 5.6.

Table 5.2. PLSR results for predicting glucose and sucrose for sliced samples and whole tubers using selected wavelengths obtained by IPLS and GA (from sampled wavelengths) and VIS/NIR interactance in the 2009 and 2011 seasons for Frito Lay 1879 and Russet Norkotah cultivars.

Season	Wavelength selection technique	Sample type	Cultivar _{constituent}	Preprocessing ^a	Window Width	Calibration model				Validation model		
						R _{cal}	RMSEC	RMSEC _{cv} (%)	LVs	R _{pred}	RMSEP (%)	RPD
2009	IPLS	Slice	FL _{GL}	A ₀ ,B ₁ ,C ₀	W1	0.79	0.0166	0.0177	8	0.78	0.0162	1.61
			FL _{SU}	A ₀ ,B ₁ ,C ₂	W1	0.42	0.0726	0.0764	19	0.40	0.0760	1.09
			RN _{GL}	A ₇ ,B ₃ ,C ₀	W1	0.97	0.0174	0.0228	16	0.95	0.0247	3.02
			RN _{SU}	A ₀ ,B ₁ ,C ₀	W1	0.69	0.0572	0.0692	20	0.64	0.0621	1.30
		Whole	FL _{GL}	A ₀ ,B ₁ ,C ₀	W1	0.83	0.0150	0.0169	20	0.81	0.0151	1.72
			FL _{SU}	A ₀ ,B ₁ ,C ₂	W1	0.51	0.7292	0.7793	20	0.43	0.7772	1.10
			RN _{GL}	A ₀ ,B ₁ ,C ₀	W1	0.98	0.0129	0.0189	20	0.97	0.0192	3.89
			RN _{SU}	A ₀ ,B ₁ ,C ₂	W1	0.63	0.0941	0.1103	19	0.51	0.1052	1.15
	GA	Slice	FL _{GL}	A ₉ ,B ₁ ,C ₀		0.79	0.0174	0.0183	13	0.78	0.0175	1.59
			FL _{SU}	A ₉ ,B ₁ ,C ₀		0.36	0.0234	0.0239	5	0.34	0.0247	1.07
			RN _{GL}	A ₇ ,B ₁ ,C ₀		0.95	0.0215	0.0237	13	0.94	0.0263	2.83
			RN _{SU}	A ₀ ,B ₁ ,C ₀		0.61	0.0626	0.0722	17	0.49	0.0703	1.15
		Whole	FL _{GL}	A ₁ ,B ₁ ,C ₀		0.84	0.0151	0.0183	20	0.80	0.0167	1.66
			FL _{SU}	A ₇ ,B ₁ ,C ₁		0.52	0.7235	0.7975	20	0.43	0.7715	1.10
			RN _{GL}	A ₀ ,B ₁ ,C ₀		0.98	0.0121	0.0209	20	0.97	0.0204	3.66
			RN _{SU}	A ₉ ,B ₁ ,C ₀		0.71	0.0557	0.0733	20	0.55	0.0684	1.18
2011	IPLS	Slice	FL _{GL}	A ₀ ,B ₁ ,C ₀	W2	0.80	0.0281	0.0374	19	0.68	0.0362	1.30
			FL _{SU}	A ₇ ,B ₁ ,C ₀	W1	0.74	0.0470	0.0571	20	0.71	0.0436	1.43
			RN _{GL}	A ₇ ,B ₂ ,C ₂	W1	0.94	0.0326	0.0652	20	0.88	0.0497	2.12
			RN _{SU}	A ₄ ,B ₂ ,C ₂	W3	0.81	0.1235	0.2131	20	0.78	0.1267	1.57
		Whole	FL _{GL}	A ₀ ,B ₂ ,C ₀	W3	0.91	0.0213	0.0426	20	0.71	0.0397	1.32
			FL _{SU}	A ₀ ,B ₃ ,C ₀	W2	0.84	0.0320	0.0472	19	0.80	0.0384	1.64
			RN _{GL}	A ₀ ,B ₁ ,C ₂	W1	0.95	0.0251	0.0625	20	0.91	0.0453	2.08
			RN _{SU}	A ₀ ,B ₁ ,C ₀	W1	0.95	0.0830	0.1565	6	0.94	0.1081	2.82
	GA	Slice	FL _{GL}	A ₇ ,B ₃ ,C ₀		0.78	0.0337	0.0414	20	0.74	0.0363	1.49
			FL _{SU}	A ₇ ,B ₁ ,C ₀		0.89	0.0285	0.0440	20	0.81	0.0391	1.70
			RN _{GL}	A ₅ ,B ₃ ,C ₂		0.92	0.0391	0.0635	20	0.84	0.0516	1.87
			RN _{SU}	A ₇ ,B ₁ ,C ₂		0.61	0.1656	0.2028	6	0.41	0.1843	1.08
		Whole	FL _{GL}	A ₄ ,B ₁ ,C ₁		0.48	0.4802	0.5320	6	0.41	0.4673	1.11
			FL _{SU}	A ₆ ,B ₁ ,C ₁		0.22	0.7480	0.7593	1	0.21	0.6494	1.03
			RN _{GL}	A ₀ ,B ₁ ,C ₂		0.95	0.0257	0.0696	13	0.71	0.0672	1.40
			RN _{SU}	A ₀ ,B ₃ ,C ₀		0.82	0.1325	0.2612	5	0.77	0.1916	1.56

^a See table 3.1 footnote.



Nomenclature: R_{VC}

R= correlation coefficient of prediction model

V= Cultivar (Frito Lay 1879 (FL), or Russet Norkotah (RN)).

C= Constituent (glucose (GL), or sucrose (SU)).

Figure 5.6. Best relationships between wet chemistry based and PLSR predicted constituents for Frito Lay 1879 and Russet Norkotah in the 2009 season for a) Glucose for sliced samples, b) Glucose for whole tubers, c) Sucrose for sliced samples, and d) Sucrose for whole tubers.

Table 5.3 shows the number of selected wavelengths from VIS/NIR interactance data for potato tubers in the case of both sugars, for FL and RN cultivars, and for the 2009 and 2011 seasons. GA produces more selected wavelengths than IPLS in both visual and near-infrared

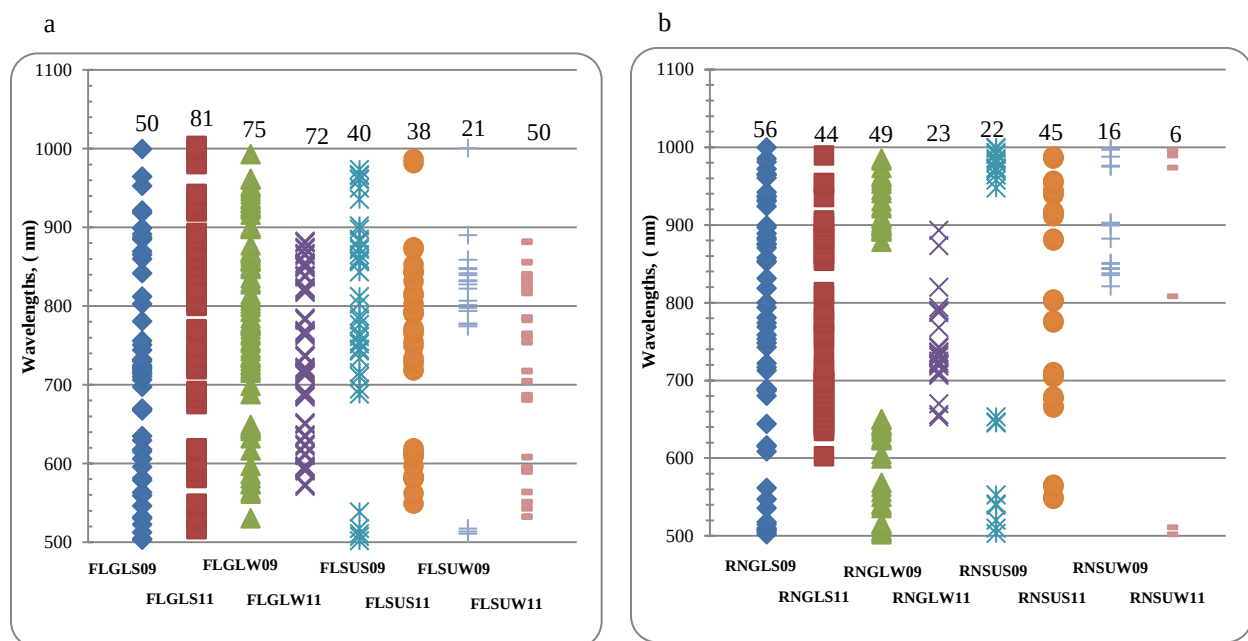
regions which explains the better performance of prediction models based on GA than IPLS. However, in practical applications, and in the case of closer performance between the two methods, IPLS is preferred as it needs less time for prediction than GA.

Table 5.3. Selected wavelengths for predicting glucose and sucrose for sliced samples and whole tubers using IPLS and GA methods (from sampled wavelengths) and VIS/NIR interactance in the 2009 and 2011 seasons for Frito Lay 1879 and Russet Norkotah cultivars.

Season	Wavelength selection technique	Sample type	Cultivar _{constituent}	Window width	Total no. of wavelengths	No. of wavelengths in VIS region	No. of wavelengths in NIR region
2009	IPLS	Slice	FL _{GL}	W=1	29	15	14
			FL _{SU}	W=1	20	11	9
			RN _{GL}	W=1	56	27	29
			RN _{SU}	W=1	68	32	36
		Whole	FL _{GL}	W=1	75	35	40
			FL _{SU}	W=1	21	5	16
			RN _{GL}	W=1	49	27	22
			RN _{SU}	W=1	33	18	15
	GA	Slice	FL _{GL}	W=1	165	108	57
			FL _{SU}	W=1	202	120	82
			RN _{GL}	W=1	202	116	86
			RN _{SU}	W=1	165	116	49
		Whole	FL _{GL}	W=1	184	97	87
			FL _{SU}	W=1	193	116	77
			RN _{GL}	W=1	217	118	99
			RN _{SU}	W=1	182	94	88
2011	IPLS	Slice	FL _{GL}	W=2	11	11	0
			FL _{SU}	W=1	19	19	0
			RN _{GL}	W=1	14	14	0
			RN _{SU}	W=3	18	16	2
		Whole	FL _{GL}	W=3	24	18	6
			FL _{SU}	W=2	20	18	2
			RN _{GL}	W=1	37	35	2
			RN _{SU}	W=1	24	18	6
	GA	Slice	FL _{GL}	W=1	229	133	96
			FL _{SU}	W=1	247	149	98
			RN _{GL}	W=1	239	136	103
			RN _{SU}	W=1	228	130	102
		Whole	FL _{GL}	W=1	214	146	68
			FL _{SU}	W=1	228	138	90
			RN _{GL}	W=1	207	123	84
			RN _{SU}	W=1	229	136	93

As can be noted from Fig 5.6, there were common wavelengths between the two seasons. Due to the different number of samples used in each season, variation between samples, and more storage conditions used in the 2011 season, there were some differences in the number of selected wavelengths. In general, the selected wavelengths in the 2011 season seemed more

efficient in yielding prediction models. With further studies, it is feasible to test the selected wavelengths and evaluate the efficacy of them to produce more robust results.



Nomenclature: VCSY

V= Cultivar (Frito Lay 1879 (FL), or Russet Norkotah (RN)).

C= Constituent (glucose (GL), or sucrose (SU)).

S= Sample type (slice (S), or whole tube (W)).

Y= Season (2009 (09), or 2011 (11)).

Figure 5.7. Schematic representation of the selected wavelengths, using VIS/NIR interactance mode and IPLS, associated with the best PLSR models of glucose and sucrose in the 2009 and 2011 seasons for sliced samples and whole tubers for a) Frito Lay 1879, b) Russet Norkotah.

5.3.4 Artificial Neural Network (ANN) Results

5.3.4.1 Full and sampled variables models

Results for prediction models of glucose and sucrose for FL and RN using different types of artificial neural network for full and sampled wavelengths for are shown in table 5.4. For models based on full wavelengths, sliced samples in the 2009 season showed slightly less performance than PLSR for full wavelengths. Values of R(RPD) for glucose prediction models were 0.89(2.24) for FL using RBFNN and 0.86(1.91) for RN using FFNN. For sucrose models,

the values were 0.58(1.15) for FL using NEWGRNN and 0.27(0.97) for RN using FFNN. For whole tubers, correlation performance was close to PLSR results with R(RPD) values of 0.77(1.44) for FL using RBFNNE and 0.95(3.09) for RN in the case of glucose using RBFNN. For the sucrose models, R(RPD) values were 0.46(1.01) for FL obtained using RBFNN and 0.63(1.28) for RN using RBFNNE. Results for the 2011 season showed improvement in both sugars' correlation for sliced samples. Glucose prediction models showed R(RPD) values as of 0.92(2.35) for FL obtained using RBFNNE and 0.94(2.97) for RN obtained using RBFNN. Sucrose prediction models showed R(RPD) values as high as 0.82(1.67) and 0.36(1.08) for RN using FFNN in both cases. Models for whole samples or whole tubers showed weak correlation performance.

Results of prediction models obtained using sampled wavelengths mostly showed the same results as the same models using full wavelengths for both seasons using ANN. As an exception, in the 2009 season, an improvement in RN sucrose prediction resulted for sliced samples with R(RPD) values as of 0.52(1.15) using NEWGRNN. Similar performance to the 2701 (full) wavelengths was achieved for both glucose and sucrose in the case of whole tubers. In the 2011 season, also the same performance was achieved for glucose and sucrose prediction models in the case of both cultivars. Consequently, using sampled wavelengths yielded same performance for both seasons and it reduced computation time.

Table 5.4. ANN results for predicting glucose and sucrose for sliced samples and whole tubers using VIS/NIR interactance and using full (2701) and sampled wavelengths in the 2009 and 2011 seasons for Frito Lay 1879 and Russet Norkotah cultivars.

Season	Wavelengths utilized	Sample type	Cultivar _{Constituent}	ANN type, characteristics	Training			Testing		
					R _{train}	SeCV _{train} (%)	RPD	R _{test}	SeP _{test} (%)	RPD
2009	Full	Slice	FL _{GL}	RBFNN	0.89	0.01	2.16	0.89	0.0121	2.24
			FL _{SU}	NEWGRNN	0.65	0.02	1.38	0.58	0.0199	1.15
			RN _{GL}	FFNN, 1500	0.91	0.03	2.17	0.86	0.0324	1.91
			RN _{SU}	FFNN, 1000	0.36	0.07	1.02	0.27	0.0721	0.97
		Whole	FL _{GL}	RBFNNE	0.78	0.0172	1.55	0.77	0.0182	1.44
			FL _{SU}	RBFNN	0.55	0.02	1.13	0.46	0.0225	1.01
			RN _{GL}	RBFNN	0.96	0.0224	2.51	0.95	0.0208	3.09
			RN _{SU}	RBFNNE	0.75	0.06	1.21	0.63	0.0547	1.28
	Sampled	Slice	FL _{GL}	RBFNN	0.90	0.0120	2.28	0.90	0.0118	2.29
			FL _{SU}	RBFNN	0.64	0.0189	1.35	0.57	0.0204	1.12
			RN _{GL}	FFNN, 1000	0.89	0.0343	1.95	0.86	0.0315	1.96
			RN _{SU}	NEWGRNN	0.58	0.0636	1.15	0.52	0.0608	1.15
		Whole	FL _{GL}	RBFNN	0.77	0.0181	1.56	0.76	0.0189	1.44
			FL _{SU}	RBFNNE	0.55	0.0226	1.13	0.45	0.0226	1.01
			RN _{GL}	RBFNN	0.97	0.0178	3.34	0.94	0.0226	2.89
			RN _{SU}	RBFNNE	0.75	0.0603	1.21	0.63	0.0547	1.28
2011	Full	Slice	FL _{GL}	RBFNNE	0.93	0.0206	2.70	0.92	0.0222	2.35
			FL _{SU}	FFNN, 500	0.84	0.0359	1.70	0.82	0.0394	1.67
			RN _{GL}	RBFNN	0.97	0.0157	3.31	0.94	0.0170	2.97
			RN _{SU}	FFNN, 1500	0.34	0.2221	1.71	0.36	0.4323	1.08
		Whole	FL _{GL}	FFNN, 1500	0.42	0.0465	1.11	0.36	0.0510	1.07
			FL _{SU}	FFNN, 1000	0.35	0.0523	0.99	0.28	0.0498	1.01
			RN _{GL}	FFNN, 500	0.37	0.2365	0.88	0.29	0.1878	0.83
			RN _{SU}	FFNN, 1000	0.39	0.1698	1.06	0.31	0.2381	1.05
	Sampled	Slice	FL _{GL}	RBFNN	0.92	0.0222	2.51	0.90	0.0244	2.14
			FL _{SU}	RBFNN	0.96	0.0182	3.34	0.95	0.0200	3.29
			RN _{GL}	FFNN, 1000	0.56	0.1567	1.22	0.37	0.1895	1.06
			RN _{SU}	FFNN, 1500	0.17	0.4167	0.80	0.14	0.5356	0.84
		Whole	FL _{GL}	FFNN, 1000	0.36	0.0593	0.87	0.33	0.0515	1.06
			FL _{SU}	RBFNN	0.96	0.0150	3.45	0.94	0.0168	3.01
			RN _{GL}	NEWGRNN	0.46	0.2092	1.09	0.17	0.2412	0.63
			RN _{SU}	FFNN, 1500	0.51	0.2170	1.14	0.45	0.3221	1.11

5.3.4.2 Selected variables- ANN models

Results of ANN prediction models based on variable selection techniques, IPLS and GA in the case of the 2009 and 2011 seasons are shown in table 5.5. In the 2009 season, for sliced samples, FL and RN glucose prediction models using IPLS showed R(RPD) values of 0.67(1.35) obtained using FFNN and 0.95(3.16) using RBFNN respectively which is slightly better, for RN, than the values obtained using full 2701 wavelengths or sampled ones. Sucrose prediction models, however, showed less performance for RN compared with full or sampled wavelengths models with R(RPD) values of 0.56(1.09) using RBFNN and 0.20(0.99) using FFNN for FL and RN respectively.

Whole tubers' prediction models using selected wavelengths showed almost the same performance for glucose prediction compared to full or sampled wavelengths models with R(RPD) values of 0.77(1.49) for FL and 0.95(3.21) for RN using RBFNN in both cases. Sucrose prediction models for FL and RN, however, showed less correlation statistics, for RN, compared to those for full or sampled wavelengths models with R(RPD) values, obtained from GA, of 0.46(1.01) and 0.16(0.99) using RBFNN and NEWGRNN respectively. Results also showed that both IPLS and GA resulted in similar performance for glucose and sucrose prediction models in the case of both cultivars in 2009 season.

For selected wavelengths prediction models in the 2011 season using ANN as shown in table 5.5, generally considerably lower correlation was obtained compared with full or samples wavelengths models in contrast to the results achieved in the case of PLSR with an exception of the glucose prediction model for sliced samples for FL in which R(RPD) values were 0.91(2.25) obtained using GA and RBFNN. Such results give priority to the PLSR prediction method over

ANN for the application of variable selection on achieving the same or even better efficiency in predicting glucose and sucrose for potato tubers using the VIS/NIR interactance technique.

Table 5.5. ANN results for predicting glucose and sucrose for sliced samples and whole tubers using selected wavelengths obtained by IPLS and GA (from sampled wavelengths) and VIS/NIR interactance in the 2009 and 2011 seasons for Frito Lay 1879 and Russet Norkotah cultivars.

Season	Wavelength selection technique	Sample type	Cultivar _{Constituent}	ANN type, characteristics	Training			Testing		
					R _{train}	SeCV _{train} (%)	RPD	R _{test}	Sep _{test} (%)	RPD
2009	IPLS	Slice	FL _{GL}	FFNN, 150	0.73	0.0197	1.43	0.67	0.0202	1.35
			FL _{SU}	RBFNN	0.53	0.0250	1.03	0.54	0.0227	1.00
			RN _{GL}	RBFNN	0.97	0.0177	3.79	0.95	0.0196	3.16
			RN _{SU}	FFNN, 500	0.16	0.0657	1.08	0.15	0.0812	1.00
		Whole	FL _{GL}	RBFNN	0.80	0.0163	1.64	0.77	0.0175	1.49
			FL _{SU}	RBFNN	0.55	0.0225	1.14	0.45	0.0224	1.02
			RN _{GL}	RBFNN	0.96	0.0220	2.56	0.95	0.0201	3.21
			RN _{SU}	FFNN, 1000	0.21	0.0691	1.02	0.12	0.0845	0.96
	GA	Slice	FL _{GL}	FFNN, 1000	0.72	0.0190	1.41	0.71	0.0184	1.42
			FL _{SU}	FFNN, 500	0.62	0.0201	1.28	0.56	0.0210	1.09
			RN _{GL}	RBFNN	0.96	0.0195	3.44	0.95	0.0190	3.25
			RN _{SU}	FFNN, 300	0.30	0.0728	1.00	0.20	0.0711	0.99
		Whole	FL _{GL}	RBFNN	0.78	0.0168	1.59	0.77	0.0174	1.50
			FL _{SU}	RBFNN	0.49	0.0227	1.13	0.46	0.0226	1.01
			RN _{GL}	RBFNN	0.96	0.0219	2.57	0.95	0.0210	3.06
			RN _{SU}	NEWGRNN	0.29	0.0728	1.00	0.16	0.0712	0.99
2011	IPLS	Slice	FL _{GL}	RBFNN	0.92	0.0216	2.58	0.90	0.0239	2.18
			FL _{SU}	FFNN, 500	0.34	0.0573	1.06	0.25	0.0636	1.03
			RN _{GL}	FFNN, 500	0.51	0.1588	1.15	0.48	0.1714	1.14
			RN _{SU}	FFNN, 50	0.36	0.2593	1.37	0.35	0.4327	1.05
		Whole	FL _{GL}	FFNN, 500	0.36	0.0536	0.96	0.19	0.0546	1.00
			FL _{SU}	FFNN, 1000	0.53	0.0510	1.10	0.30	0.0705	1.03
			RN _{GL}	FFNN, 50	0.65	0.1703	0.94	0.40	0.1600	1.08
			RN _{SU}	FFNN, 1000	0.58	0.2044	1.08	0.50	0.2572	1.10
	GA	Slice	FL _{GL}	RBFNN	0.92	0.0222	2.50	0.91	0.0231	2.25
			FL _{SU}	FFNN, 50	0.24	0.0596	1.02	0.22	0.06	1.02
			RN _{GL}	FFNN, 150	0.44	0.1735	1.06	0.30	0.1918	1.02
			RN _{SU}	FFNN, 50	0.36	0.2360	1.61	0.21	0.4769	0.98
		Whole	FL _{GL}	FFNN, 300	0.23	0.3252	1.03	0.16	0.4466	1.01
			FL _{SU}	FFNN, 150	0.40	0.0557	0.93	0.34	0.0475	1.06
			RN _{GL}	FFNN, 300	0.33	0.1876	1.10	0.24	0.1658	0.94
			RN _{SU}	FFNN, 50	0.53	0.2072	1.20	0.54	0.3008	1.19

5.3.5 Results of Potatoes Classification Based on Sugar Levels and Selected Wavelengths

Based on glucose and sucrose thresholds as described in section 5.3.3, data was divided into two classes for sliced samples and whole tubers for the 2009 and 2011 seasons as presented in table 5.6 with outliers removed. Classification error for training, using cross validation, and testing groups for both seasons is shown in table 5.7 with the lowest classification error in each case marked with bold font.

Table 5.6. Number of samples in each class based on glucose and sucrose levels, obtained from wet chemistry, for sliced samples and whole tubers in the 2009 and 2011 seasons for Frito Lay 1879 and Russet Norkotah cultivars.

Season	Sample	Cultivar _{Constituent}	Class 1 (less than threshold)	Class 2 (above threshold)
2009	Slice	FL_{GL}	445	562
		FL_{SU}	523	458
		RN_{GL}	177	159
		RN_{SU}	195	139
	Whole	FL_{GL}	222	281
		FL_{SU}	266	229
		RN_{GL}	88	79
		RN_{SU}	87	68
2011	Slice	FL_{GL}	204	186
		FL_{SU}	218	146
		RN_{GL}	66	84
		RN_{SU}	58	82
	Whole	FL_{GL}	136	57
		FL_{SU}	122	71
		RN_{GL}	31	44
		RN_{SU}	26	48

In both seasons, classification performance generally followed the PLSR trend explained in section 5.3.3. For the 2009 season, classification error values of glucose-based models (16% and 13% for FL and RN in the case of sliced samples, and 18%, and 13% in the case of whole tubers) were much lower than those for sucrose-based models (35%, and 36% for FL and RN in the case of sliced samples and 34%, and 38% for FL and RN in the case of whole tubers). However, results for 2011 indicated better performance for sucrose-based classification. Classification errors based on glucose were 21%, and 23% for FL and RN in the case of sliced samples and 23%, and 0% for FL and RN in the case of whole tubers. While for sucrose-based models, error values were 23%, and 18% for FL and RN in the case of slice samples and 26%,

and 14% for FL and RN in the case of whole tubers. Results for 2011 showed the advantage of obtaining broader sugar distribution, especially for sucrose, which was confirmed by PLSR prediction outputs.

As noted from table 5.7, PLSDA generally presented the least classification error, especially for the 2011 season, followed by LDA models. Knn and ANN, however, did not yield as powerful performance as the former methods. Additionally, classifier fusion models showed similar results to PLSDA in many cases, with the lowest error obtained for RN using glucose levels in the 2011 season for whole tubers. Consequently, combining classifier outputs did improve classification results in certain cases.

Classification results obtained in this study show the potential of sorting potato tubers based on glucose or sucrose levels associated with, and of importance to, processing for each sugar, which has not been addressed before using any non-destructive method. Such sorting is important for the frying industry and can help decrease the losses during storage by identifying tubers with excessive sugar levels such that the possibility exists for reversing sugar levels to normal levels using the recommended temperatures for a certain period in a process called reconditioning (Sowokinos, 2007).

Table 5.7. Classification results of sliced samples and whole tubers of Frito Lay 1879 and Russet Norkotah cultivars based on glucose and sucrose levels and using multiple classification techniques and VIS/NIR interactance in the 2009 and 2011 seasons.

Season	Sample type	Cultivar ^{Constituent}	Preprocessing for LDA; Knn; PLSDA; ANN; combined classifier ^a	Training error (%)				Testing error (%)				
				LDA	Knn	PLSDA	ANN	LDA	Knn	PLSDA	ANN	Combined classifiers
2009	Slice	FL_{GL}	A ₇ ; A ₉ ; A ₆ ; A ₄ ;A ₄	19	22	16	21	17	22	16	20	16
		FL_{SU}	A ₄ ; A ₀ ; A ₇ ; A ₆ ;A ₇	36	43	35	38	35	44	38	41	38
		RN_{GL}	A ₇ ; A ₉ ; A ₅ ; A ₇ ;A ₅	16	19	13	18	15	24	13	15	13
		RN_{SU}	A ₉ ; A ₅ ; A ₀ ; A ₀ ;A ₀	34	41	26	40	42	36	41	38	41
	Whole	FL_{GL}	A ₀ ; A ₆ ; A ₁₂ ; A ₄ ;A ₀	21	25	13	28	21	24	18	25	19
		FL_{SU}	A ₉ ; A ₄ ; A ₇ ; A ₉ ;A ₄	36	44	35	45	34	35	35	47	35
		RN_{GL}	A ₇ ; A ₉ ; A ₅ ; A ₀ ;A ₉	17	25	7	19	13	28	18	51	18
		RN_{SU}	A ₁₂ ; A ₀ ; A ₄ ; A ₀ ;A ₄	28	39	24	40	44	41	38	38	38
2011	Slice	FL_{GL}	A ₇ ; A ₇ ; A ₇ ; A ₇ ;A ₇	24	42	16	36	31	32	21	33	21
		FL_{SU}	A ₀ ; A ₁₂ ; A ₆ ; A ₆ ;A ₀	29	43	2	30	38	40	23	40	33
		RN_{GL}	A ₀ ; A ₀ ; A ₆ ; A ₀ ;A ₆	25	41	2	33	30	40	23	40	23
		RN_{SU}	A ₆ ; A ₀ ; A ₁₂ ; A ₀ ;A ₅	23	34	18	29	43	39	18	32	18
	Whole	FL_{GL}	A ₄ ; A ₀ ; A ₅ ; A ₀ ;A ₅	28	44	22	23	28	29	23	29	23
		FL_{SU}	A ₄ ; A ₄ ; A ₇ ; A ₆ ;A ₆	25	41	26	26	39	29	26	34	32
		RN_{GL}	A ₅ ; A ₀ ; A ₉ ; A ₄ ;A ₄	9	23	0	29	21	39	7	29	0
		RN_{SU}	A ₀ ; A ₁₂ ; A ₄ ; A ₀ ;A ₄	7	30	0	14	36	29	14	21	14

^a See table 3.1 footnote.

5.3.6 Results for 2009-2011 Combined Data

Results of combining data from the 2009 and 2011 seasons for building prediction models for glucose and sucrose using PLSR are shown in table 5.8. Results obtained of both sugars for FL and RN showed similar results compared to those obtained from the 2009 season in the case of sliced samples and whole tubers. Consequently, combining data from both seasons didn't show significant improvement for prediction models compared to results conducted from the 2009 season using PLSR.

Table 5.8. PLSR results for predicting glucose and sucrose for sliced samples and whole tubers using VIS/NIR interactance for Frito Lay 1879 and Russet Norkotah cultivars using 2009 and 2011 combined data.

Sample type	Cultivar _{constituent}	Preprocessing ^a	Calibration model				Validation model		
			R _{cal}	RMSEC	RMSEC _{cv}	LVs	R _{pred}	RMSEP	RPD
Slice	FL _{GL}	A ₀ ,B ₁ ,C ₁	0.79	0.4499	0.4855	20	0.78	0.4638	1.58
	FL _{SU}	A ₁₂ ,B ₃ ,C ₀	0.70	0.0325	0.0378	20	0.54	0.0408	1.18
	RN _{GL}	A ₇ ,B ₁ ,C ₂	0.98	0.0447	0.0564	20	0.96	0.0584	3.82
	RN _{SU}	A ₇ ,B ₃ ,C ₀	0.77	0.1354	0.2305	1	0.44	0.1940	1.03
Whole	FL _{GL}	A ₀ ,B ₁ ,C ₁	0.83	0.3980	0.4809	20	0.78	0.4412	1.58
	FL _{SU}	A ₀ ,B ₁ ,C ₀	0.55	0.0306	0.0318	10	0.51	0.0321	1.16
	RN _{GL}	A ₀ ,B ₁ ,C ₂	0.98	0.0377	0.0563	18	0.96	0.0571	3.42
	RN _{SU}	A ₀ ,B ₃ ,C ₀	0.47	0.1372	0.1443	3	0.46	0.1584	1.13

^a See table 3.1 footnote.

ANN prediction results obtained from combined data of 2009 and 2011 seasons are shown in table 5.9. No improvement in glucose and sucrose prediction performance was observed for either cultivar which, in addition to the previous PLSR results, gives a note that combining data from the two seasons is negatively affected by the variation in samples, and reference (glucose and sucrose) distribution that was affected by adding another storing temperature (1°C) to the experiments in the 2011 season.

Table 5.9. ANN results for predicting glucose and sucrose for sliced samples and whole tubers using VIS/NIR interactance for Frito Lay 1879 and Russet Norkotah cultivars using 2009 and 2011 combined data.

Sample Type	Cultivar _{Constituent}	ANN type, characteristics	Training			Testing		
			R _{train}	SeCV _{train} (%)	RPD	R _{test}	Sep _{test} (%)	RPD
Slice	FL _{GL}	FFNN, 1500	0.71	0.0296	1.34	0.69	0.0302	1.38
	FL _{SU}	FFNN, 1000	0.46	0.0420	1.06	0.45	0.0403	1.12
	RN _{GL}	FFNN, 500	0.87	0.1246	2.01	0.85	0.1294	1.85
	RN _{SU}	FFNN, 500	0.46	0.1625	1.68	0.40	0.1597	0.83
Whole	FL _{GL}	NEWGRNN	0.61	0.0294	1.44	0.58	0.0326	1.22
	FL _{SU}	FFNN, 500	0.31	0.0424	0.84	0.26	0.0357	0.96
	RN _{GL}	RBFNNE	0.75	0.1219	1.78	0.72	0.1527	1.36
	RN _{SU}	FFNN, 500	0.54	0.1172	1.32	0.46	0.1380	1.13

5.4 Conclusion

VIS/NIR interactance signals in the range of 501-1004 nm of potato sliced samples and whole tubers were extracted from the original wavelengths range (446-1125 nm) and used to build prediction models using partial least squares regression and different types of artificial neural network for glucose and sucrose sugars. IPLS and GA as wavelength selection techniques were applied on a sampled set of signals acquired from the VIS/NIR interactance measurements (446-1125 nm) for Frito Lay 1879 and Russet Norkotah potato cultivars. All electronic measurements were compared against glucose and sucrose that were measured using the enzymatic approach. PLSR and ANN were used to build calibration and prediction models for glucose and sucrose in the case of 0.5'' (12.7 mm) sliced samples and whole tubers. Selected wavelengths were found to have strong correlation performance with RMSEP of 0.0162%, and 0.0247% for FL and RN for sliced samples in the case of glucose. In the case of sucrose, the best models had RMSEP values of 0.0227% and 0.0621% for FL and RN respectively. Whole tubers yielded even better performance than sliced samples with RMSEP values of 0.0151, and 0.0192% for FL and RN in the case of glucose, while those values for sucrose were 0.0241% and 0.1052% for FL and RN. Such levels of accuracy are suitable for monitoring sugar levels especially for whole tubers which is crucial practice during storage, and prior to processing.

Classification of tubers based on sugar levels important to the frying industry was shown to have feasible application for sorting, especially in the case of glucose in which the error values for testing sets were as low as 18%, and 0% for FL, and RN, and those values were 26%, and 14% for sucrose. Classification performance can likely be improved with broader and more uniform distribution of sugars, and scanning the whole tuber in more than one point on the tuber surface so that more robust prediction and classification is feasible. Moreover, to simulate real sorting conditions, it is important to conduct more experiments on moving tubers mixed with clods, and using tubers that have soil attached to their surfaces.

CHAPTER 6 RAPID EVALUATION OF PHYSIOLOGICAL STATUS OF POTATO TUBERS USING NEAR-INFRARED REFLECTANCE SPECTROSCOPIC METHODS

(Expanded from: Rady, A.M., Guyer, D.E. 2014. Evaluation of sugar content in potatoes using nir reflectance and wavelength selection techniques. Postharvest Biology and Technology (in review))

6.1 Introduction

Near-infrared (NIR) reflectance is the most extensively-studied phenomenon explained using physics laws for the interaction between light and matter in the NIR region (Dahm and Dahm, 2001; Olinger et al., 2001). When NIR light interacts with a biological object, a portion of the light is reflected from the surface, yet holding limited information about the chemical composition of the object. Another portion of the incident light, however, penetrates the surface, scatters, is adsorbed by different molecules, is transmitted through the object to the other side, and/or is reflected again from the surface and holding significant of information of the object components which is known as the diffuse reflected light. Diffuse reflectance observations have been studied and many mathematical models were developed in attempts to model it starting from Lambert law. It is also strongly affected by the general radiation transfer equation.

In general, NIR diffuse reflectance became the base for most commercially-built NIR instrumentations (Shenk et al., 2001). In the agriculture and food industry fields, NIR diffuse reflectance was applied by Gera and Norris (1968), to rapidly detect moisture and protein for grains, and protein, oil, and moisture content for soybeans. Later, Shenk et al. (1977a, 1977b), studied the application of NIR diffuse reflectance on forage quality. Since then, the investigation of applying NIR diffuse reflectance techniques on monitoring quality parameters for fruits and vegetables has continued.

Potato tuber, as a major crop around the world, with its importance for human diet, and with numerous industries that aim to provide high quality fresh or processed products, results in the need for rapid, yet accurate tools by which quality characteristics monitoring could be conducted either on line for the raw tubers during harvesting, sorting, storage, and/or even after processing. Sugar content in potato tubers is very critical in determining the suitability for processing as French fry or chip, so the establishment of a device to easily, accurately and cost effectively monitor sugar levels is needed and beneficial for growers to estimate best time for harvest, and for quality control specialists in processing plants to confirm the status of tubers. In the current chapter, three data sets were obtained from the NIR reflectance system; full wavelengths, sampled wavelengths, and selected wavelengths using IPLS and GA. The analysis of such data sets included building prediction models for glucose and sucrose, and classification of sliced samples and whole tubers using various techniques.

Based on the above noted considerations, the objectives of this study were:

1. Determine the potential of NIR diffuse reflectance spectroscopy for quantifying the level of sucrose and glucose levels in potato tubers.
2. Development of a reliable prediction models that may be used to detect the amount of sucrose and glucose in potato tubers at levels which are important for potato growers and processors.
3. Detect the most effective wavelengths related to glucose and sucrose absorption based on NIR diffuse reflectance measurements and variable selection techniques.
4. Study the potential of using NIR reflectance measurements of potatoes along with wavelength selection techniques to classify whole tubers and 0.5'' (12.7 mm) sliced samples according to sugar levels related to the frying process.

6.2 Material and Methods

6.2.1 Constituent Measurement

6.2.1.1 Potato sample preparation

The sample preparation technique for the reflectance measurements followed the same routine as illustrated in section 4.2.2.1. Additionally, whole tubers were also used in electronic measurements. Whole tubers were placed on the sample holder such that the middle area of the longitudinal axis was penetrated by incident light.

6.2.1.2 Wet chemistry basis measurements

The procedure used in juice extraction from sliced samples was the same as was used in section 5.2.2.1. Also, wet chemistry steps conducted to evaluate glucose and sucrose concentrations were the same as mentioned in section 4.2.2.2 and using equations 3.2 and 3.3 respectively.

6.2.2 NIR Reflectance System

According to Burn and Ciurczak (2001), the use of NIR diffuse reflection for quantitative analysis of biological products is widely applicable. In the reflectance mode, the incident light penetrates the sample surface and a portion of such light passes within the sample tissue and is then reflected back, known as diffuse reflectance, and detected with information about the internal composition of the tubers (Chen, 1978). The light source probe tip and the detector tip were approximately 3 cm from the sample upper surface. . An InGaAs spectrometer (model No. NIR512L-1.7T1, Control Development, Inc., South Bend, IN, USA) with spectral resolution of 3.25 nm FWHM and linear dispersion of 1.625 nm/pixel was used in the reflectance mode, in the

wavelength range of 900-1685 nm along with an Oriel radiometric power supply with a 300 watt maximum power (model No.68931, Oriel Inst., Irvine, CA, USA), and an Oriel light source (model No. 66881, Oriel Inst., Irvine, CA, USA) having 250 watt maximum power, and with a quartz tungsten halogen lamp. In the diffuse reflectance experiments, the sample area covered by the light source had a diameter of 25.5 mm. The integration time was set as 4 ms, and each measurement is the average of four individual measurements. The incident light was directed on the middle area of the cut side of the slice. For the whole tubers, the light was directed on the surface approximately in the center area where the longitudinal, and perpendicular axes intersect. The detector covers an area on the sample surface of 12.7 mm diameter. A schematic diagram of the reflectance system used in the experiment is represented in Fig. 6.1. The relative reflectance was calculated using equation 3.1.

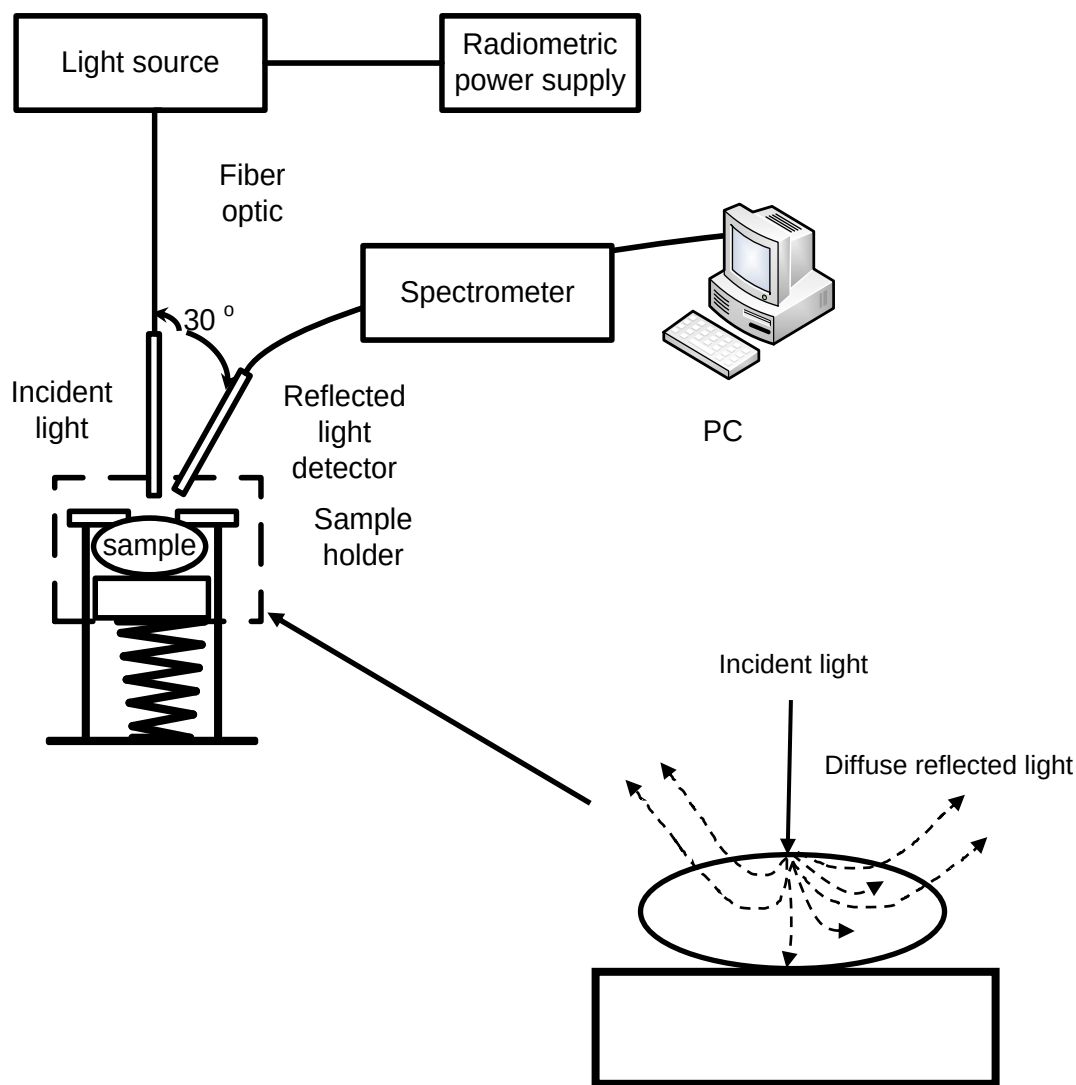


Figure.6.1. Schematic representation of NIR diffuse reflectance mode and a clearer view of sample setting.

6.2.3 Data analysis Discussion and Approach

6.2.3.1 Data handling

Various scenarios of handling and consequently statistical analysis were applied for the relative reflectance data extracted from reflectance experiments. Fig. 6.2 shows the sequence of data handling and methods used to build prediction models. First the signals resulted from measurements were visually checked for noise and consequently no reduction on number of

wavelengths was conducted. Next, two modes of data were tested, the data containing full 784 wavelengths, and data sampled at every 3 wavelengths resulting in a 262- variable matrix. Additionally, data from the 2009 and 2011 seasons were mixed together in both the full and selected wavelengths models and PLSR and ANN analysis were conducted to obtain prediction models for both glucose and sucrose on all data sets.

6.2.3.2 Partial least squares regression (PLSR)

A complete description of PLSR used in this research along with pretreatment for either spectra or reference values is described in section 3.2.4.

6.2.3.3 Artificial neural network (ANN)

The same artificial neural network types, and configuration applied in this study were the same as that used in section 4.2.4.3.

6.2.3.4 Wavelength selection

A complete description of interval partial least squares (IPLS) and genetic algorithm (GA) used in this research was listed in section 4.2.4.5.

6.2.3.5 Classification of potatoes based on sugar levels

A complete description of the techniques used in potatoes classification based on selected wavelengths is stated in section 5.2.4.5.

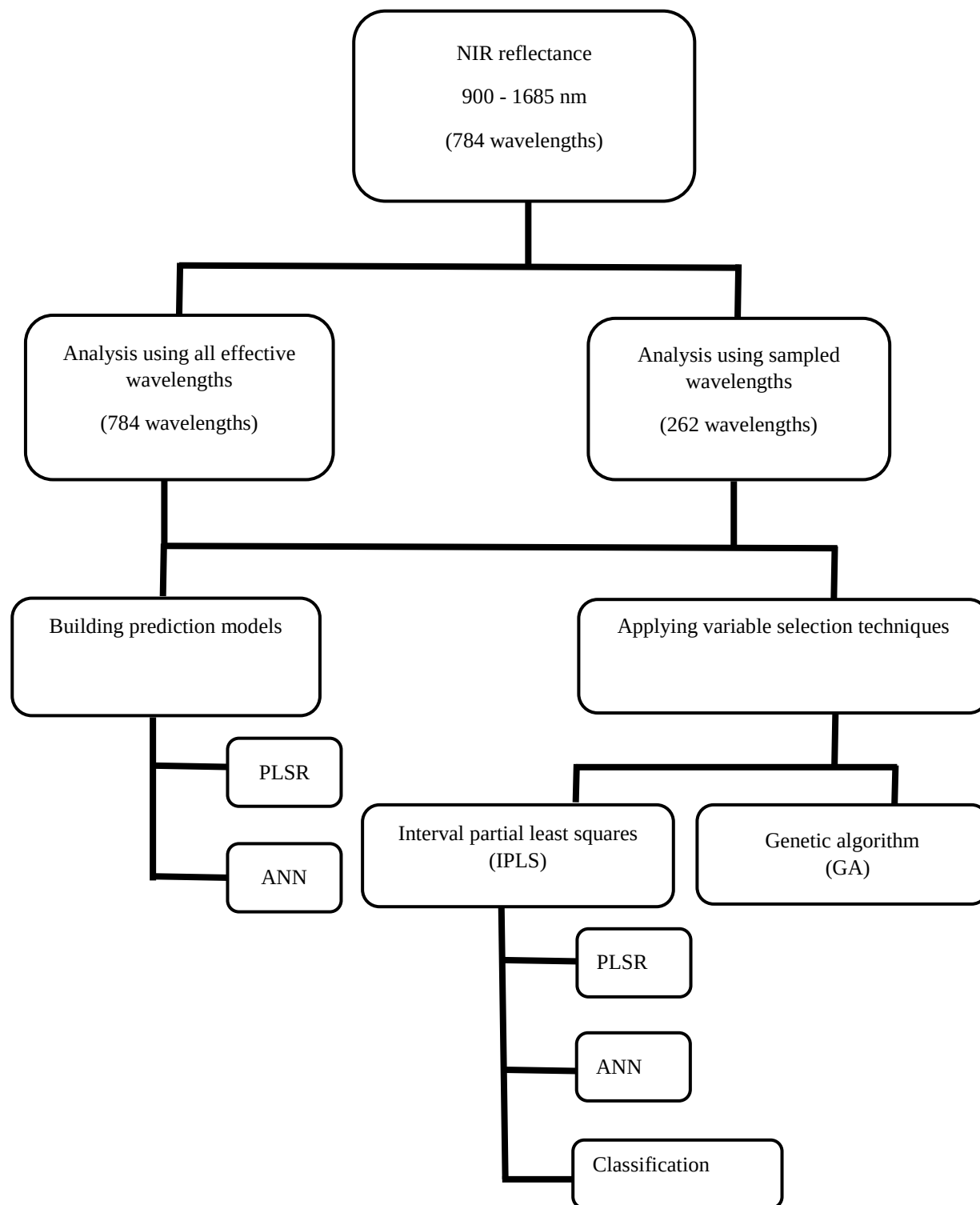


Figure 6.2. A schematic diagram of data handling and analysis for NIR reflectance experiments to obtain prediction models of glucose and sucrose for two potato cultivars.

6.3 Results and Discussions

6.3.1 Constituents Distribution

The basic statistics for both glucose and sucrose over the 2009 and 2011 data were shown in table 4.1.

6.3.2 Spectra for Sliced Samples and Whole Tubers

The relative values of absorbance or $\log(1/\text{reflectance})$ resulting from the NIR experiments through the 2009 and 2011 seasons are shown in figures 6.3 and 6.4 respectively for both Frito Lay 1879 and Russet Norkotah cultivars in the case of sliced samples and whole tubers. Signals of whole tubers in both seasons appear to be more condensed than signals of sliced samples especially in the 2011 season with an exception of RN for the whole tubers. Russet Norkotah showed more spread in signals than in Frito Lay1879 and the possible explanation for this is that the FL periderm is thinner and easier to get scraped, in the case of whole tuber measurement, than in RN giving the chance to sometimes expose the periderm layer, which is different in color, to the incident light than the outer layer thus yielding different reflectance values. Whereas in the case of RN, the outer layer is stronger, more difficult to get scraped during handling, and consequently, the variability in surface reflectance is less. Moreover, for the sliced samples, similar trend of signals for FL and RN was observed with multiple water absorption peaks around 970, 1200, and 1450 nm (Workman and Weyer, 2008). Additionally, another absorption peak, in both cultivars, is noted at 1530 in the 2009 sliced data nm which is a possible indication of an OH polymeric group located in starch (Workman and Weyer, 2008). In general, signals collected from whole tubers showed less absorption than sliced samples especially for RN which yielded more condensed response than FL. The possible explanation for such result is the thicker periderm layer of RN that reduces the captured diffuse

reflectance signals. It is worth stating that the absorption peaks for sliced samples (around 970, 1200, and 1450 nm) were noted in the whole tubers in both cultivars with more clarification in FL. Another effect of the thick skin for RN was the absence of the absorption peak at 1530 nm and the relatively slight presence, compared to sliced samples, of the same peak in FL samples in the 2009 data.

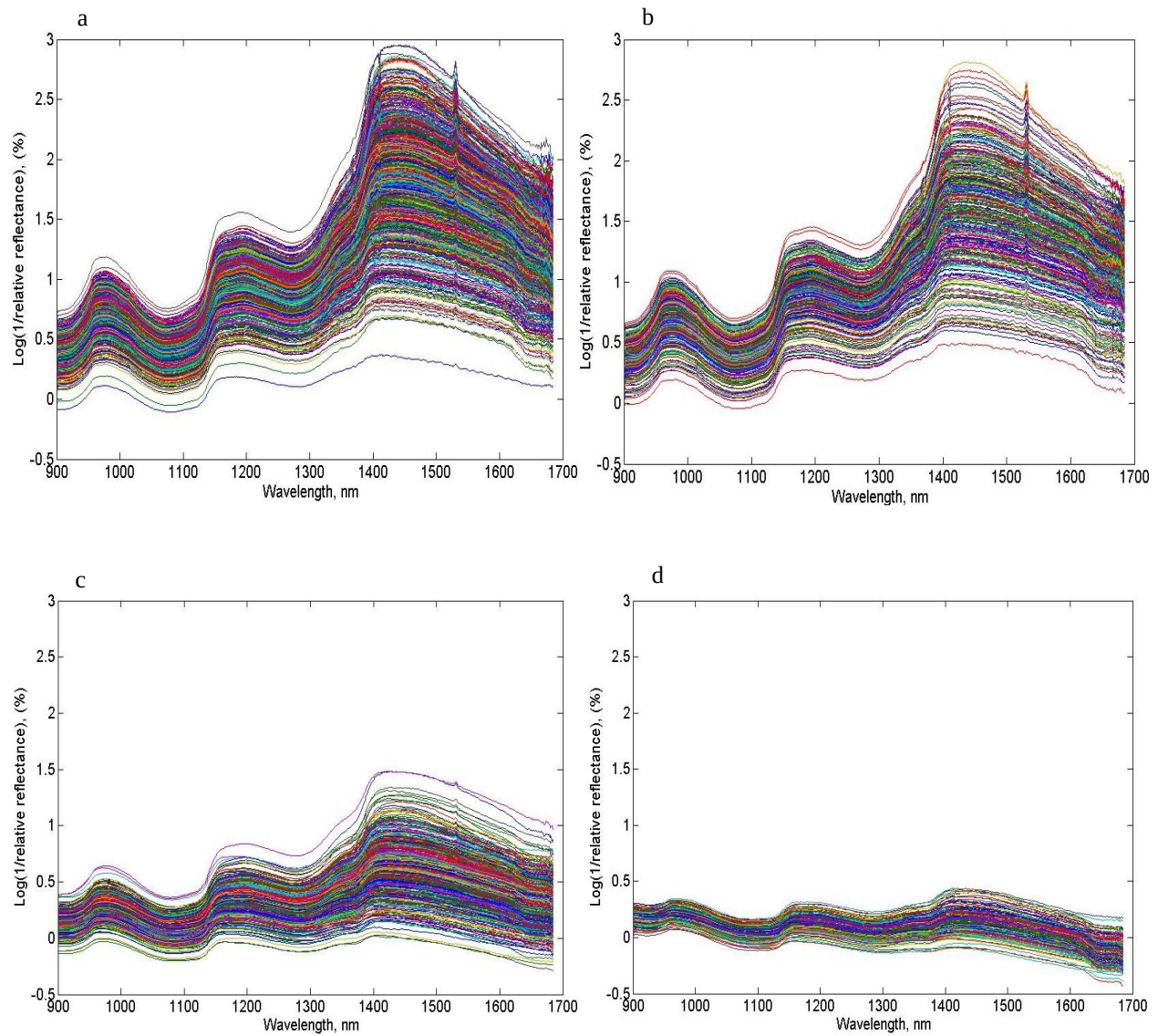


Figure 6.3. Signals of absorbance ($\log(1/\text{relative reflectance})$) for the 2009 season data for sliced samples a. Frito Lay 1879, b. Russet Norkotah, and for whole tubers for c. Frito Lay 1879, and d. Russet Norkotah.

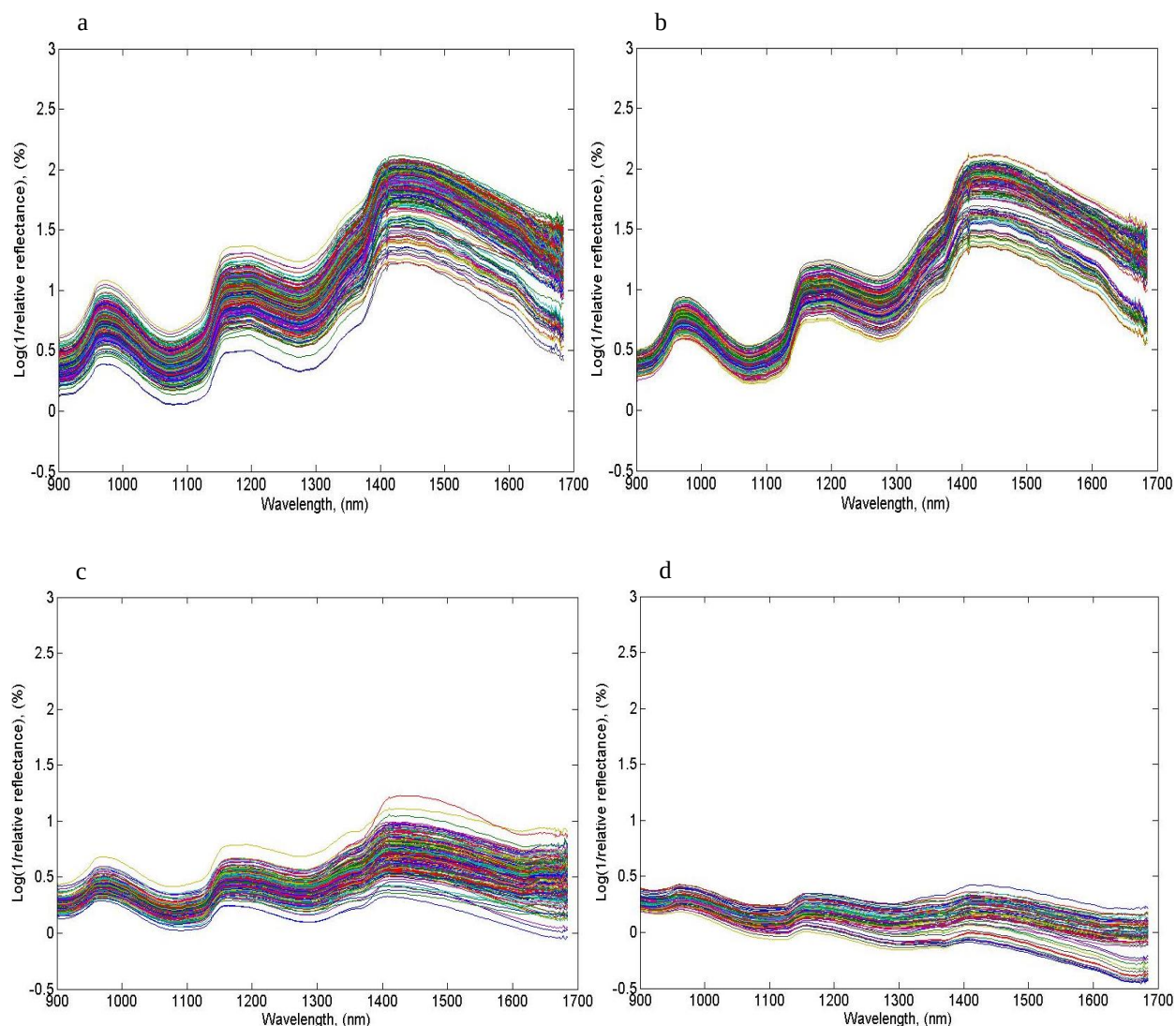


Figure 6.4. Signals of absorbance ($\log(1/\text{relative reflectance})$) for the 2011 season data for sliced samples a. Frito Lay 1879, b. Russet Norkotah, and for whole tubers for c. Frito Lay 1879, and d. Russet Norkotah.

6.3.3 Partial Least Squares Regression (PLSR) Results

6.3.3.1 Full and sampled variables models

Results for calibration and prediction models of glucose and sucrose using full wavelengths for both Frito Lay1879 (FL) and Russet Norkotah (RN) cultivars in the case of the 2009 and 2011 seasons are shown in table 6.1. In the 2009 season, and based on full wavelengths

models, for glucose prediction models, FL yielded strong correlation with R(RPD) values of 0.74(1.47) for sliced samples and 0.76(1.53) for whole tubers. Compared to glucose models, weaker correlation was obtained for sucrose with R(RPD) values of 0.36(1.06) for sliced samples and 0.40(1.05) for whole tubers. In the case of RN, better glucose prediction was obtained than with FL. Sliced samples yielded glucose prediction models with R(RPD) values of as high as 0.95(3.11) and those values were 0.98(4.24) for whole tubers. Whereas, again weaker correlation was obtained for sucrose prediction models with R(RPD) values of 0.65(1.31) for sliced samples, and 0.57(1.18) for whole tubers. For the prediction models obtained from sampled wavelengths, slightly less performance, compared to the full wavelengths' models, was obtained for both cultivars in the case of both glucose and sucrose which clarifies the advantage of reducing data dimension using sampling.

In the 2011 season, using the full wavelength range showed general higher correlation than in the 2009 season. Glucose prediction models showed R(RPD) values as high as 0.83(1.78) for sliced samples and 0.71(1.28) for whole tubers. Sucrose prediction models showed R(RPD) values of 0.61(1.26) for sliced samples and 0.65(1.33) for whole tubers. Higher prediction of sugars, than FL, was obtained for RN with R(RPD) of glucose models as of 0.97(4.21) for sliced samples and 0.98(4.84) for whole tubers. In the case of sucrose, R(RPD) values were 0.55(1.18) for sliced samples, and 0.75(1.52) for whole tubers.

Performance of prediction models for glucose models based on sampled wavelengths was similar to full wavelengths models in both cultivars and both sliced samples and whole tubers except in the case of sliced samples for FL in which lower correlation was obtained. Sucrose prediction models also showed similar correlation for FL compared to full wavelengths' models. However, RN showed slightly less correlation compared with full wavelengths' models.

Table 6.1. PLSR results for predicting glucose and sucrose for sliced samples and whole tubers using NIR reflectance and using full (784) and sampled wavelengths in the 2009 and 2011 seasons for Frito Lay 1879 and Russet Norkotah cultivars.

Season	Wavelengths utilized	Sample type	Cultivar _{constituent}	Preprocessing ^a	Calibration model				Validation model		
					R _{cal}	RMSEC	RMSEC _{cv} (%)	LVs	R _{pred}	RMSEP (%)	RPD
2009	Full	Slice	FL _{GL}	A ₀ ,B ₁ ,C ₀	0.77	0.0476	0.0533	20	0.74	0.0509	1.47
			FL _{SU}	A ₅ ,B ₁ ,C ₀	0.51	0.0290	0.0311	18	0.36	0.0311	1.06
			RN _{GL}	A ₁₂ ,B ₃ ,C ₂	0.97	0.0321	0.0550	16	0.95	0.0419	3.11
			RN _{SU}	A ₁₂ ,B ₃ ,C ₂	0.82	0.0708	0.1155	9	0.65	0.0922	1.31
		Whole	FL _{GL}	A ₁₂ ,B ₃ ,C ₂	0.80	0.0446	0.0525	20	0.76	0.0492	1.53
			FL _{SU}	A ₁₂ ,B ₃ ,C ₀	0.64	0.0242	0.0286	20	0.40	0.0298	1.05
			RN _{GL}	A ₇ ,B ₃ ,C ₀	0.99	0.0109	0.0428	19	0.98	0.0228	4.24
			RN _{SU}	A ₄ ,B ₃ ,C ₀	0.77	0.0524	0.0732	3	0.57	0.0687	1.18
	Sampled	Slice	FL _{GL}	A ₁₂ ,B ₃ ,C ₀	0.77	0.0480	0.0538	18	0.73	0.0512	1.46
			FL _{SU}	A ₇ ,B ₁ ,C ₀	0.49	0.0274	0.0293	17	0.37	0.0294	1.07
			RN _{GL}	A ₁₂ ,B ₃ ,C ₂	0.96	0.0374	0.0560	17	0.93	0.0468	2.78
			RN _{SU}	A ₁₂ ,B ₁ ,C ₂	0.67	0.0614	0.0776	19	0.55	0.0687	1.18
		Whole	FL _{GL}	A ₇ ,B ₃ ,C ₀	0.74	0.0195	0.0242	19	0.71	0.0204	1.41
			FL _{SU}	A ₁₂ ,B ₃ ,C ₀	0.58	0.0255	0.0292	20	0.41	0.0291	1.08
			RN _{GL}	A ₄ ,B ₃ ,C ₂	0.99	0.0152	0.0590	9	0.96	0.0356	3.69
			RN _{SU}	A ₇ ,B ₃ ,C ₀	0.76	0.0537	0.0731	4	0.57	0.0690	1.18
2011	Full	Slice	FL _{GL}	A ₁₂ ,B ₃ ,C ₂	0.91	0.0319	0.0568	20	0.83	0.0435	1.78
			FL _{SU}	A ₆ ,B ₃ ,C ₀	0.62	0.0544	0.0630	12	0.61	0.0494	1.26
			RN _{GL}	A ₄ ,B ₁ ,C ₀	0.99	0.0295	0.0973	20	0.97	0.0468	4.21
			RN _{SU}	A ₁₂ ,B ₁ ,C ₀	0.67	0.0614	0.0776	19	0.55	0.0687	1.18
		Whole	FL _{GL}	A ₇ ,B ₃ ,C ₀	0.97	0.0146	0.0452	19	0.71	0.0441	1.28
			FL _{SU}	A ₅ ,B ₃ ,C ₀	0.67	0.0437	0.0526	12	0.65	0.0475	1.33
			RN _{GL}	A ₀ ,B ₁ ,C ₀	0.99	0.0290	0.0881	18	0.98	0.0387	4.84
			RN _{SU}	A ₉ ,B ₁ ,C ₂	0.76	0.1060	0.1700	7	0.75	0.1324	1.52
	Sampled	Slice	FL _{GL}	A ₄ ,B ₁ ,C ₂	0.77	0.0503	0.0616	15	0.72	0.0534	1.45
			FL _{SU}	A ₅ ,B ₁ ,C ₀	0.67	0.0478	0.0531	20	0.59	0.0484	1.20
			RN _{GL}	A ₄ ,B ₃ ,C ₁	0.99	0.0655	0.1855	20	0.97	0.0976	4.20
			RN _{SU}	A ₀ ,B ₁ ,C ₂	0.71	0.1477	0.1963	8	0.38	0.2049	0.97
		Whole	FL _{GL}	A ₁₂ ,B ₁ ,C ₀	0.91	0.0245	0.0465	19	0.71	0.0445	1.27
			FL _{SU}	A ₆ ,B ₃ ,C ₀	0.63	0.0454	0.0543	10	0.62	0.0491	1.29
			RN _{GL}	A ₄ ,B ₁ ,C ₀	0.99	0.0310	0.1017	17	0.98	0.0402	4.65
			RN _{SU}	A ₅ ,B ₃ ,C ₂	0.87	0.0809	0.1836	15	0.67	0.1459	1.38

^a See table 3.1 footnote.

6.3.3.2 Selected variables- PLSR models

Prediction models for glucose and sucrose in the case of Frito Lay 1879 and Russet Norkotah using PLSR based on selected wavelengths from IPLS and GA for the 2009 and 2011 seasons are shown in tables 6.2 and the clarification of selected wavelengths is shown in table 6.3. In general, IPLS-based prediction models yielded better performance than GA-based models. Thus, the results for IPLS will be stated here. Glucose prediction models for sliced samples yielded R(RPD) values of 0.76(1.54) for FL and 0.94(2.73) for RN. The values for whole tubers were 0.72(1.44) for FL, and 0.95(3.05) for RN. In the case of sucrose prediction models for sliced samples, R(RPD) values were 0.50(1.15) for FL, and 0.35(1.04) for RN. The values for whole tubers were 0.45(1.12) for FL, and 0.56(1.19) for RN. By comparing the results obtained from selected wavelengths to those obtained from full wavelengths, some notes should be listed. In general, IPLS is preferable over GA as it yielded higher correlation and less selected wavelengths. Performance of prediction models obtained from IPLS was closer to, or better than, full models for both sliced samples or whole tubers with an exception of the sucrose model of RN in the case of sliced samples in which significantly lower performance was achieved. For the 2011 season, a significant improvement was achieved in the prediction performance for both sugars in the case of FL and RN and for both sliced samples and whole tubers with an exception being the glucose prediction model for FL in the case of sliced samples. Again, only results obtained using IPLS will be noted here as they showed better performance than GA models. Glucose prediction models for sliced samples showed R(RPD) values of 0.74(1.48) for FL and 0.97(4.07) for RN. The values for whole tubers were 0.82(1.78) for FL and 0.98(4.57) for RN. In the case of sucrose prediction models for sliced samples, R(RPD) values were 0.74(1.41) for FL

and 0.81(1.66) for RN. For the whole tubers, R(RPD) values were 0.73(1.46) for FL and 0.93(2.77) for RN.

Most results obtained from IPLS, table 6.2, were based on window width of two ($w=2$). GA selected variables were all with window width of one. As mentioned before, IPLS showed less number of selected variables compared to GA for 2009 and 2011 as shown in table 6.3. Moreover, all of the selected wavelengths' ranges showed a domination of the wavelengths in the range 900-1160 nm which supports that the effective wavelengths in the NIR region associated with high correlation is located within this range. The best relationships between the measured, and predicted sugar values for FL, and RN for sliced samples and whole tubers in the 2011 season is shown in Fig. 6.4.

The improvement of results for whole tubers compared to sliced samples, especially for sucrose models in the 2011 season, is possibly a result of the sugar distribution inside tubers. According to Kumar and Ezekiel (2004); and Rastovski et al. (1987), sugars inside potato tubers tend to concentrate more on the vascular ring than on other tuber parts. Consequently, the diffuse reflected light is expected to hold information of the tissue closer to the skin than to the pith.

Results also showed that prediction models of glucose and sucrose obtained using IPLS yielded better performance than GA models for both cultivars, and also table 6.3 indicates that the selected wavelengths using IPLS were less than GA in all models. Possible reasons for such results include the more likelihood for over fitting to occur in GA than IPLS in the case of fewer number of samples than variables which was noted in the 2011 season compared to 2009 season (Wise et al., 2006). Due to the lower sugar concentration for FL than RN, results showed less correlation of prediction models for FL than RN in the case of glucose and sucrose as the

detection of certain chemical substance using spectroscopic systems increase with the concentration.

It is worth stating that the prediction of glucose and sucrose for potatoes using selected wavelengths and NIR diffuse reflectance was not previously published and the prediction results obtained in this study by PLSR are comparable with others reported by Hartmann and Büning-Pfaue (1998), on homogenized samples (RMSEP= 0.041% and 0.037% for glucose and sucrose); Yaptenco et al. (2000), on whole tubers (RMSEP= 0.087% and 1.473% for glucose and sucrose); or Haase (2011), on aliquots samples (SEP=0.0389%, and 0.0966% for reducing sugars and sucrose). Sampling times in this study are significantly lower than that for all previous studies except for Yaptenco et al. (2000), which did not include a separate prediction data set. This study is also confirms the results obtained by Rady et al. (2014), in which a potential investigation of measuring glucose and sucrose of potatoes was shown using different techniques and strong correlation for glucose was achieved for sliced samples (RMSEP= 0.0515%, and 0.0786% for FL, and RN), and whole tubers (RMSEP= 0.0620%, and 0.1529% for FL, and RN) using VIS/NIR interactance spectroscopy. However, no variable selection was applied.

Table 6.2. PLSR results for predicting glucose and sucrose for sliced samples and whole tubers using selected wavelengths obtained by IPLS and GA (from sampled wavelengths) and NIR reflectance in the 2009 and 2011 seasons for Frito Lay 1879 and Russet Norkotah cultivars.

Season	Variable selection technique	Sample type	Cultivar ^{constituent}	Preprocessing ^a	Window Width	Calibration model				Validation model		
						R _{cal}	RMSEC (%)	RMSEC _{cv} (%)	LVs	R _{pred}	RMSEP (%)	RPD
2009	IPLS	Slice	FL _{GL}	A ₀ ,B ₁ ,C ₂	W1	0.78	0.0466	0.0497	20	0.76	0.0489	1.54
			FL _{SU}	A ₇ ,B ₁ ,C ₀	W2	0.58	0.0277	0.0306	19	0.50	0.0286	1.15
			RN _{GL}	A ₀ ,B ₁ ,C ₀	W1	0.96	0.0229	0.0332	20	0.94	0.0353	2.73
			RN _{SU}	A ₅ ,B ₁ ,C ₁	W1	0.54	0.7831	0.9260	20	0.35	0.9150	1.04
		Whole	FL _{GL}	A ₀ ,B ₃ ,C ₂	W2	0.73	0.0520	0.0553	20	0.72	0.0524	1.44
			FL _{SU}	A ₀ ,B ₁ ,C ₀	W2	0.52	0.0289	0.0305	20	0.45	0.0295	1.12
			RN _{GL}	A ₇ ,B ₃ ,C ₂	W2	0.96	0.0360	0.0471	20	0.95	0.0432	3.05
			RN _{SU}	A ₀ ,B ₁ ,C ₀	W1	0.69	0.0593	0.0665	20	0.56	0.0681	1.19
	GA	Slice	FL _{GL}	A ₁₂ ,B ₃ ,C ₂		0.71	0.0536	0.0596	9	0.70	0.0541	1.39
			FL _{SU}	A ₆ ,B ₁ ,C ₀		0.44	0.0303	0.0319	13	0.36	0.0307	1.07
			RN _{GL}	A ₇ ,B ₁ ,C ₂		0.93	0.0442	0.0609	13	0.87	0.0644	2.04
			RN _{SU}	A ₆ ,B ₁ ,C ₂		0.44	0.1119	0.1238	11	0.26	0.1179	0.99
		Whole	FL _{GL}	A ₄ ,B ₃ ,C ₂		0.75	0.0499	0.0592	18	0.71	0.0530	1.42
			FL _{SU}	A ₇ ,B ₁ ,C ₀		0.51	0.0290	0.0324	18	0.36	0.0312	1.06
			RN _{GL}	A ₇ ,B ₃ ,C ₀		0.97	0.0264	0.0542	18	0.94	0.0326	2.85
			RN _{SU}	A ₄ ,B ₃ ,C ₀		0.77	0.0529	0.0734	3	0.57	0.0690	1.18
2011	IPLS	Slice	FL _{GL}	A ₇ ,B ₃ ,C ₀	W3	0.86	0.0310	0.0425	17	0.74	0.0403	1.48
			FL _{SU}	A ₇ ,B ₁ ,C ₀	W1	0.90	0.0277	0.0344	20	0.74	0.0411	1.41
			RN _{GL}	A ₀ ,B ₀ ,C ₀	W2	0.98	0.0400	0.0601	20	0.97	0.0483	4.07
			RN _{SU}	A ₀ ,B ₃ ,C ₀	W2	0.98	0.0645	0.2791	20	0.81	0.2296	1.66
		Whole	FL _{GL}	A ₆ ,B ₁ ,C ₀	W1	0.87	0.0281	0.0387	19	0.82	0.0318	1.78
			FL _{SU}	A ₀ ,B ₁ ,C ₀	W2	0.87	0.0246	0.0385	20	0.73	0.0359	1.46
			RN _{GL}	A ₀ ,B ₁ ,C ₀	W2	0.99	0.0312	0.0674	20	0.98	0.0409	4.57
			RN _{SU}	A ₁₂ ,B ₁ ,C ₀	W3	0.96	0.0688	0.2085	20	0.93	0.1128	2.77
	GA	Slice	FL _{GL}	A ₉ ,B ₁ ,C ₂		0.82	0.0447	0.0583	19	0.77	0.0491	0.82
			FL _{SU}	A ₀ ,B ₁ ,C ₀		0.76	0.0452	0.0607	19	0.56	0.0554	1.12
			RN _{GL}	A ₇ ,B ₁ ,C ₁		0.89	0.1951	0.2762	11	0.87	0.2013	2.03
			RN _{SU}	A ₉ ,B ₁ ,C ₁		0.65	0.9061	1.1459	6	0.44	0.9785	1.09
		Whole	FL _{GL}	A ₇ ,B ₁ ,C ₀		0.74	0.0389	0.0481	10	0.72	0.0386	1.46
			FL _{SU}	A ₀ ,B ₁ ,C ₀		0.66	0.0440	0.0533	10	0.57	0.0513	1.23
			RN _{GL}	A ₀ ,B ₁ ,C ₀		0.96	0.0547	0.1045	13	0.93	0.0692	2.70
			RN _{SU}	A ₅ ,B ₁ ,C ₀		0.81	0.1578	0.2148	16	0.77	0.1836	1.52

^a See table 3.1 footnote.

Table 6.3. Selected wavelengths for predicting glucose and sucrose for sliced samples and whole tubers using IPLS and GA methods (from sampled wavelengths) and NIR reflectance in the 2009 and 2011 seasons for Frito Lay 1879 and Russet Norkotah cultivars.

Season	Wavelength selection technique	Sample type	Cultivar ^{constituent}	Window width	Total no. of wavelengths	Minimum value (nm)	Maximum value (nm)
2009	IPLS	Slice	FL _{GL}	W=1	66	903	1160
			FL _{SU}	W=1	20	900	1156
			RN _{GL}	W=1	67	900	1131
			RN _{SU}	W=2	12	942	1111
		Whole	FL _{GL}	W=1	66	900	1157
			FL _{SU}	W=2	56	916	1159
			RN _{GL}	W=1	58	902	1153
			RN _{SU}	W=1	12	992	1157
	GA	Slice	FL _{GL}	W=1	161	900	1158
			FL _{SU}	W=1	147	902	1161
			RN _{GL}	W=1	176	901	1161
			RN _{SU}	W=1	147	900	1160
		Whole	FL _{GL}	W=1	171	900	1159
			FL _{SU}	W=1	136	900	1160
			RN _{GL}	W=1	182	900	1160
			RN _{SU}	W=1	151	900	1157
2011	IPLS	Slice	FL _{GL}	W=1	27	900	1159
			FL _{SU}	W=1	21	900	1143
			RN _{GL}	W=2	54	904	1157
			RN _{SU}	W=2	14	906	1155
		Whole	FL _{GL}	W=1	16	900	1063
			FL _{SU}	W=1	16	900	1156
			RN _{GL}	W=1	30	900	1161
			RN _{SU}	W=1	8	900	910
	GA	Slice	FL _{GL}	W=1	136	904	1160
			FL _{SU}	W=1	148	902	1161
			RN _{GL}	W=1	153	900	1157
			RN _{SU}	W=1	140	900	1160
		Whole	FL _{GL}	W=1	120	900	1161
			FL _{SU}	W=1	132	901	1156
			RN _{GL}	W=1	27	900	1159
			RN _{SU}	W=1	21	900	1143

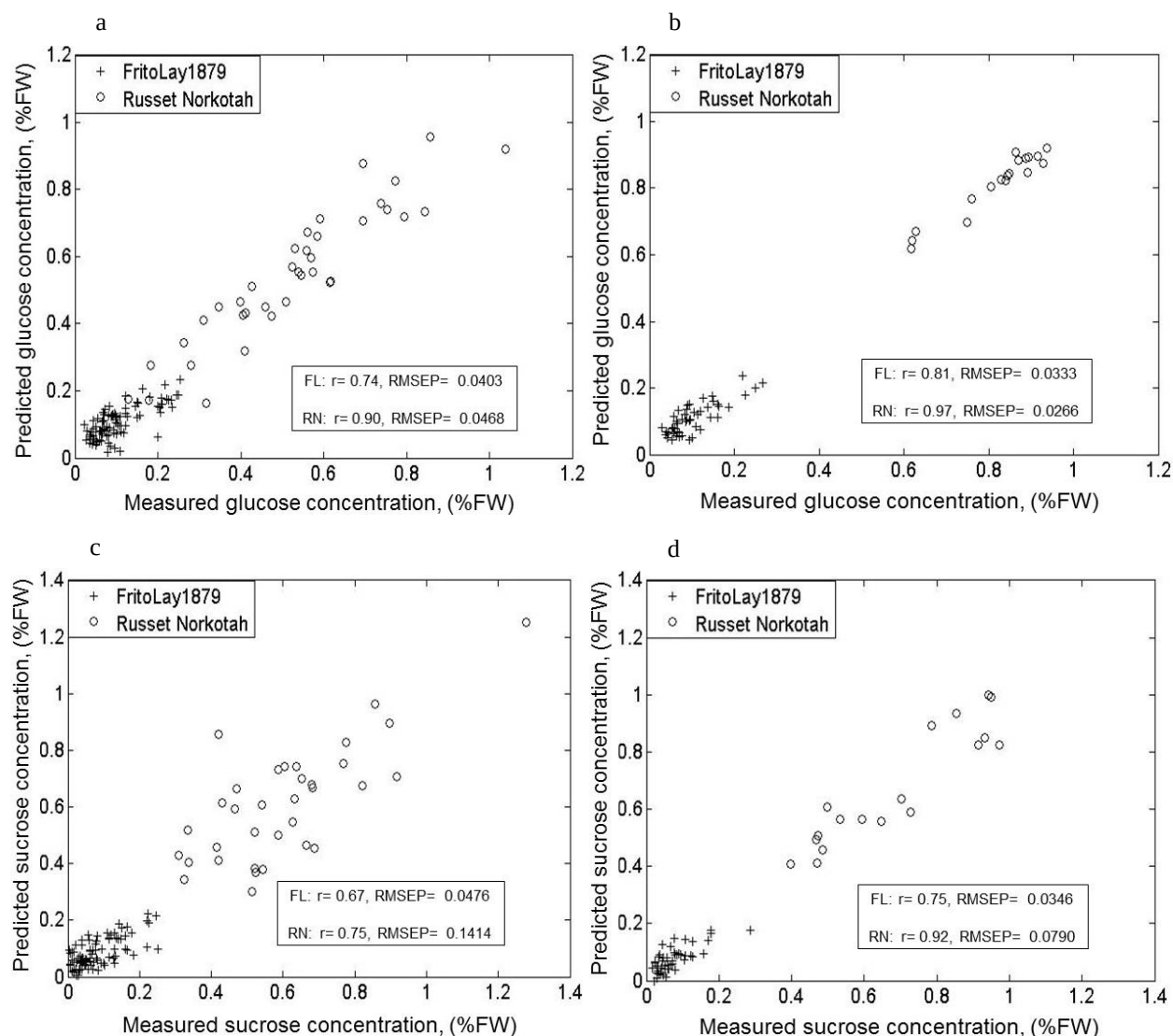


Figure 6.4. Best prediction models based on selected wavelengths and PLSR predicted constituents in the 2011 season for Frito Lay 1879 and Russet Norkotah cultivars for a) Glucose for sliced samples, b) Glucose for whole tubers, c) Sucrose for sliced samples, and d) Sucrose for whole tubers.

6.3.4 Artificial Neural Network (ANN) Results

6.3.4.1 Full and sampled variables models

Results for prediction models of glucose and sucrose for Frito Lay1879 and Russet Norkotah using different types of artificial neural network and full and sampled wavelengths in

the case of the 2009 and 2011 seasons are shown in table 6.4. In the 2009 season, sliced samples glucose prediction models showed R(RPD) values as high as 0.96(3.47) for FL and 0.93(2.86) for RN. Whole tubers showed values of 0.75(1.42) for FL and 0.95(3.29) for RN. Sucrose prediction models of sliced samples showed R(RPD) values of 0.72(1.30) for FL and 0.96(3.73) for RN. However, the values of whole tubers were 0.68(1.31) for FL and 0.68(1.32) for RN. Prediction models obtained from 2011 for sliced samples showed lower performance than 2009 season except in the case of sucrose model for FL that showed significant improvement with R(RPD) values of 0.95(3.39). For the whole tubers, lower performance was obtained for glucose prediction models. However, sucrose models showed an improvement for both cultivars with R(RPD) values of 0.96(3.80) for FL and 0.97(3.78) for RN. Most of the prediction models in table 6.4 were obtained using either RBFNN or RBFNNE except with two models in which FFNN was implemented.

Results of using ANN in models based on sampled wavelengths in the 2009 season for sliced samples showed similar correlation performance for glucose prediction models compared to full wavelength models except in the case of sucrose prediction models for FL that showed less correlation statistics. In the case of whole tubers, similar correlation was obtained for both sugars and cultivars compared to results based on all variables. In the case of the 2011 season, sliced samples showed improvement of glucose prediction for FL with R(RPD) values of 0.69(1.35). For whole tubers, glucose prediction models showed similar correlation to full-wavelength models, whereas sucrose prediction models showed better performance than the full-wavelength models.

Table 6.4. ANN results for predicting glucose and sucrose for sliced samples and whole tubers using NIR reflectance and using full (784) and sampled wavelengths in the 2009 and 2011 seasons for Frito Lay 1879 and Russet Norkotah cultivars.

Season	Wavelengths	Sample type	Cultivar ^{Constituent}	ANN type, characteristics	Training			Testing		
					R _{train}	SeCV _{train} (%)	RPD	R _{test}	Sep _{test} (%)	RPD
2009	All	Slice	FL _{GL}	RBFNNE	0.96	0.01	3.55	0.96	0.0086	3.47
			FL _{SU}	RBFNNE	0.78	0.02	1.42	0.72	0.0220	1.30
			RN _{GL}	RBFNNE	0.93	0.03	2.72	0.93	0.0290	2.86
			RN _{SU}	RBFNN	0.97	0.02	3.85	0.96	0.0202	3.73
		Whole	FL _{GL}	RBFNNE	0.79	0.0183	1.59	0.75	0.0201	1.42
			FL _{SU}	RBFNNE	0.70	0.02	1.46	0.68	0.0219	1.31
			RN _{GL}	RBFNN	0.97	0.02	3.94	0.95	0.0252	3.29
			RN _{SU}	RBFNNE	0.78	0.06	1.57	0.68	0.0570	1.32
	Sampled	Slice	FL _{GL}	RBFNNE	0.96	0.0078	3.70	0.95	0.0086	3.47
			FL _{SU}	RBFNNE	0.70	0.0276	1.12	0.63	0.0261	1.09
			RN _{GL}	RBFNNE	0.94	0.0330	2.88	0.93	0.0317	2.62
			RN _{SU}	RBFNN	0.97	0.0232	3.97	0.97	0.0192	3.92
		Whole	FL _{GL}	RBFNN	0.78	0.02	1.57	0.73	0.0205	1.39
			FL _{SU}	RBFNNE	0.70	0.0212	1.46	0.68	0.0219	1.31
			RN _{GL}	RBFNN	0.96	0.02	3.66	0.95	0.0213	3.64
			RN _{SU}	RBFNNE	0.78	0.0586	1.57	0.68	0.0570	1.32
2011	All	Slice	FL _{GL}	RBFNN	0.45	0.0737	0.85	0.47	0.0754	0.77
			FL _{SU}	RBFNN	0.96	0.0185	3.35	0.95	0.0180	3.39
			RN _{GL}	RBFNNE	0.74	0.1401	1.31	0.61	0.2171	0.90
			RN _{SU}	FFNN, 1000	0.40	0.2645	1.34	0.24	0.4844	0.94
		Whole	FL _{GL}	RBFNN	0.55	0.0585	0.88	0.34	0.0988	0.55
			FL _{SU}	RBFNNE	0.97	0.0142	3.64	0.96	0.0163	3.80
			RN _{GL}	FFNN, 1000	0.31	0.1556	1.05	0.27	0.2010	0.86
			RN _{SU}	RBFNNE	0.98	0.0525	4.92	0.97	0.0830	3.78
	Sampled	Slice	FL _{GL}	RBFNN	0.48	0.0741	0.85	0.37	0.0809	0.72
			FL _{SU}	RBFNN	0.96	0.0185	3.35	0.95	0.0191	3.20
			RN _{GL}	RBFNNE	0.70	0.1459	1.26	0.69	0.1446	1.35
			RN _{SU}	FFNN, 500	0.49	0.2047	1.74	0.13	0.4661	0.98
		Whole	FL _{GL}	RBFNN	0.56	0.0591	0.87	0.48	0.0610	0.89
			FL _{SU}	RBFNN	0.96	0.0158	3.28	0.95	0.0192	3.23
			RN _{GL}	RBFNN	0.71	0.1275	1.26	0.63	0.1707	1.01
			RN _{SU}	RBFNNE	0.98	0.0525	4.92	0.97	0.0830	3.78

6.3.4.2 Selected variables- ANN models

Results of ANN prediction models based on variable selection techniques, IPLS and GA, for the 2009 and 2011 seasons are shown in tables 6.5. In the 2009 season, models for sliced samples using either IPLS or GA showed close correlation performance compared to full wavelengths results for glucose prediction of RN and sucrose prediction of FL. Other models showed less correlation statistics. Whole tuber FL and RN glucose prediction models showed close correlation performance compared to full wavelengths' models.

The IPLS and GA variable selection prediction models for the 2011 season generally resulted in better performance than full-wavelength models in the case of sliced samples for glucose prediction with values of R(RPD) were 0.72(1.20) for FL and 0.75(1.20) for RN using IPLS for both models. Sucrose prediction models, however, showed similar correlation to that obtained using full wavelengths. Prediction models obtained using GA showed similar performance to the full wavelengths models with an improvement in glucose prediction for RN that resulted R(RPD) values of 0.62(0.77).

Generally, ANN results showed that both IPLS and GA resulted in similar performance for glucose and sucrose prediction models in the case of both cultivars for sliced samples and whole tubers, with some exceptions as mentioned in the case of whole tubers for 2011 season. Thus, preference for IPLS is given as it showed a general trend for fewer selected variables and comparable or better correlation as GA.

Table 6.5. ANN results for predicting glucose and sucrose for sliced samples and whole tubers using selected wavelengths obtained by IPLS and GA (from sampled wavelengths) and NIR reflectance in the 2009 and 2011 seasons for Frito Lay 1879 and Russet Norkotah cultivars.

Season	Wavelength selection technique	Sample type	Cultivar ^{Constituent}	ANN type, characteristics	Training			Testing		
					R _{train}	SeCV _{train} (%)	RPD	R _{test}	Sep _{test} (%)	RPD
2009	IPLS	Slice	FL _{GL}	FFNN, 150	0.96	0.0078	3.71	0.96	0.0086	3.47
			FL _{SU}	RBFNNE	0.67	0.0264	1.17	0.62	0.0273	1.05
			RN _{GL}	RBFNN	0.83	0.0552	1.64	0.71	0.0715	1.08
			RN _{SU}	RBFNNE	0.97	0.0229	4.02	0.96	0.0171	4.40
		Whole	FL _{GL}	RBFNN	0.79	0.0191	1.52	0.78	0.0190	1.58
			FL _{SU}	RBFNN	0.65	0.0225	1.38	0.62	0.0232	1.23
			RN _{GL}	RBFNNE	0.94	0.0358	2.80	0.90	0.0408	2.18
			RN _{SU}	RBFNNE	0.78	0.0586	1.57	0.68	0.0570	1.32
	GA	Slice	FL _{GL}	FFNN, 150	0.41	0.0268	1.08	0.36	0.0290	1.06
			FL _{SU}	RBFNN	0.54	0.0339	0.91	0.57	0.0354	0.81
			RN _{GL}	RBFNN	0.88	0.0452	2.00	0.78	0.0672	1.15
			RN _{SU}	RBFNN	0.97	0.0232	3.97	0.96	0.0183	4.11
		Whole	FL _{GL}	RBFNNE	0.78	0.0197	1.48	0.75	0.0201	1.42
			FL _{SU}	RBFNNE	0.70	0.0212	1.46	0.68	0.0219	1.31
			RN _{GL}	RBFNN	0.96	0.0244	3.70	0.96	0.0208	3.72
			RN _{SU}	RBFNNE	0.78	0.0586	1.57	0.68	0.0570	1.32
2011	IPLS	Slice	FL _{GL}	RBFNNE	0.73	0.1508	1.22	0.72	0.1626	1.20
			FL _{SU}	RBFNNE	0.97	0.0166	3.73	0.96	0.0157	3.89
			RN _{GL}	RBFNNE	0.73	0.1508	1.22	0.75	0.1626	1.20
			RN _{SU}	FFNN, 150	0.34	0.2125	1.79	0.23	0.4503	1.04
		Whole	FL _{GL}	RBFNN	0.60	0.0593	0.87	0.39	0.0883	0.62
			FL _{SU}	FFNN, 50	0.34	0.0491	1.05	0.25	0.0598	1.04
			RN _{GL}	RBFNNE	0.79	0.1434	1.12	0.55	0.1722	1.00
			RN _{SU}	RBFNNE	0.98	0.0525	4.92	0.97	0.0830	3.78
	GA	Slice	FL _{GL}	FFNN, 1000	0.28	0.0950	0.60	0.27	0.1116	0.46
			FL _{SU}	RBFNN	0.96	0.0181	3.42	0.96	0.0167	3.66
			RN _{GL}	RBFNNE	0.77	0.1308	1.40	0.68	0.1499	1.30
			RN _{SU}	FFNN, 500	0.25	0.3025	1.26	0.23	0.4591	1.02
		Whole	FL _{GL}	FFNN, 100	0.33	0.0490	1.05	0.31	0.0513	1.06
			FL _{SU}	RBFNN	0.96	0.0163	3.17	0.95	0.0189	3.28
			RN _{GL}	RBFNNE	0.64	0.2074	0.82	0.62	0.2291	0.77
			RN _{SU}	RBFNNE	0.98	0.0525	4.92	0.97	0.0830	3.78

6.3.5 Results of Potatoes Classification Based on Sugar Levels and Selected Wavelengths

The highest classification rate values of training and testing groups obtained for slice samples and whole potato tubers of FL and RN cultivars based on glucose and sucrose concentrations for both seasons are shown in table 6.6 with the number of class 1 (sugar level < the threshold) or class 2 (sugar level > the threshold). Classification performance generally followed the PLSR trend stated in section 6.3.3. For the 2009 season, classification error values of glucose-based models for the sliced samples (17% and 19% for FL and RN), were similar to values obtained for whole tubers (19%, and 17% for FL and RN). Sucrose-based classification models, however, yielded lower performance for sliced samples (37% and 32% for FL and RN), and whole tubers (38% and 31% for FL and RN). Classification results for glucose in 2011 showed similar, or lightly lower performance compared to 2009 results for the sliced samples (18% and 23% for FL and RN) and better results in the case of whole tubers especially for RN (23% and 0% for FL and RN). Moreover, significantly enhanced classification rates were obtained for sucrose models in the case of sliced samples (25% and 18% for FL and RN), and whole tubers (29%, and 21%).

In general, LDA, PLSDA, and classifier fusion yielded better classification results than other techniques (Knn, and ANN). Such trend is a result of the capability of PLSDA technique, as illustrated in section 2.4.3, for treating data with colinearity problem, and the application of PCA analysis on spectra data prior to performing classification using LDA. Combined classifiers also resulted in better classification than Knn, and ANN classifiers, and slightly similar to results obtained by PLSDA, and LDA.

Sugar distribution in the 2011 season resulted in better classification results, compared with the 2009 season, especially for sucrose which follows the same trend obtained in PLSR.

Classification of potatoes based on sugar levels and using noninvasive measurements was not addressed before and results showed the potential for classifying tubers with sugar content that is not suitable for frying such that they can potentially be reconditioned to reduce sugar content (Sowokinos, 2007). Enhancing classification outputs beyond those obtained in this study is feasible by developing broader sugar distribution, increasing the number of samples, and using kernel-based classification methods (i.e. soft independent modeling of class analogy or SIMCA, Gaussian mixture models, and support vector machines or SVM).

Table 6.6. Classification results of sliced samples and whole tubers of Frito Lay 1879 and Russet Norkotah cultivars based on glucose and sucrose levels and using multiple classification techniques and NIR reflectance in the 2009 and 2011 seasons.

Season	Sample type	Number of samples		Cultivar _{constituent}	Preprocessing for LDA; Knn; PLSDA; ANN; combined classifier ^a	Training error (%)				Testing error (%)				
		Class 1	Class 2			LDA	Knn	PLSDA	ANN	LDA	Knn	PLSDA	ANN	Combined classifiers
2009	Slice	445	445	FL _{GL}	A ₉ ; A ₅ ; A ₄ ; A ₁₂ ; A ₇	21	34	18	22	17	26	21	26	20
		523	523	FL _{SU}	A ₇ ; A ₇ ; A ₆ ; A ₉ ; A ₁₂	37	45	34	44	37	43	38	40	38
		177	177	RN _{GL}	A ₄ ; A ₄ ; A ₀ ; A ₇ ; A ₁₀	16	38	6	29	19	25	19	25	21
		195	195	RN _{SU}	A ₆ ; A ₆ ; A ₁₀ ; A ₀ ; A ₁₀	32	47	33	44	35	44	35	42	32
	Whole	222	222	FL _{GL}	A ₇ ; A ₇ ; A ₁₂ ; A ₉ ; A ₇	28	30	19	24	27	26	20	26	19
		266	266	FL _{SU}	A ₀ ; A ₅ ; A ₁₃ ; A ₀ ; A ₁₂	35	43	30	43	43	39	39	45	38
		88	88	RN _{GL}	A ₀ ; A ₅ ; A ₀ ; A ₉ ; A ₇	12	25	12	25	17	19	19	19	19
		87	87	RN _{SU}	A ₆ ; A ₀ ; A ₇ ; A ₀ ; A ₇	31	36	31	46	37	31	31	42	31
2011	Slice	204	204	FL _{GL}	A ₇ ; A ₉ ; A ₀ ; A ₉ ; A ₀	27	34	12	34	29	26	18	28	18
		218	218	FL _{SU}	A ₇ ; A ₄ ; A ₇ ; A ₀ ; A ₄	31	45	15	36	25	40	33	38	33
		66	66	RN _{GL}	A ₁₂ ; A ₉ ; A ₇ ; A ₁₂ ; A ₇	19	35	13	30	40	33	23	33	23
		58	58	RN _{SU}	A ₁₂ ; A ₄ ; A ₀ ; A ₄ ; A ₀	20	36	11	36	43	32	18	29	18
	Whole	136	136	FL _{GL}	A ₁₂ ; A ₀ ; A ₀ ; A ₀ ; A ₀	19	43	3	26	23	34	23	37	23
		122	122	FL _{SU}	A ₉ ; A ₀ ; A ₉ ; A ₉ ; A ₇	17	46	21	31	29	32	29	39	37
		31	31	RN _{GL}	A ₀ ; A ₀ ; A ₄ ; A ₉ ; A ₀	2	40	0	27	27	20	0	27	0
		26	26	RN _{SU}	A ₀ ; A ₁₂ ; A ₇ ; A ₀ ; A ₇	8	29	27	27	21	43	21	43	21

^a See table 3.1 footnote.

6.3.6 Results for 2009-2011 Combined Data

Results of combining data from the 2009 and 2011 seasons for building prediction models using NIR reflectance measurements for glucose and sucrose using PLSR are shown in table 6.7. Correlation performance was not significantly improved using combined data compared to either the 2009 or 2011 results.

Table 6.7. PLSR results for predicting glucose and sucrose for sliced samples and whole tubers using NIR reflectance for Frito Lay 1879 and Russet Norkotah cultivars using 2009 and 2011 combined data.

Sample type	Cultivar _{constituent}	Preprocessing ^a	Calibration model				Validation model		
			R _{cal}	RMSEC (%)	RMSEC _{cv} (%)	LVs	R _{pred}	RMSEP (%)	RPD
Slice	FL _{GL}	A ₀ ,B ₁ ,C ₁	0.78	0.4603	0.5298	18	0.77	0.4674	1.56
	FL _{SU}	A ₇ ,B ₁ ,C ₀	0.60	0.0374	0.0411	14	0.53	0.0415	1.18
	RN _{GL}	A ₁₀ ,B ₁ ,C ₂	0.98	0.0478	0.0763	20	0.95	0.0693	3.20
	RN _{SU}	A ₄ ,B ₁ ,C ₁	0.78	0.6808	0.9203	20	0.63	0.8173	1.27
Whole	FL _{GL}	A ₀ ,B ₁ ,C ₂	0.79	0.0512	0.0622	20	0.74	0.0571	1.48
	FL _{SU}	A ₄ ,B ₃ ,C ₀	0.70	0.0284	0.0334	19	0.55	0.0325	1.17
	RN _{GL}	A ₄ ,B ₃ ,C ₂	0.99	0.0197	0.0781	7	0.95	0.0594	3.33
	RN _{SU}	A ₀ ,B ₁ ,C ₀	0.76	0.0968	0.1238	20	0.71	0.1155	1.42

^a See table 3.1 footnote.

For ANN results obtained from combined data and shown in table 6.8, lower performance was obtained when both seasons' data was combined. Results obtained from ANN in addition to the previous PLSR results gives a note that combining data from the two seasons is negatively affected by variation in samples and reference (glucose and sucrose) distribution that was broader in 2011 than in 2009 caused by adding another storage temperature (1°C) to the experiments in the 2011 season.

Table 6.8. ANN results for predicting glucose and sucrose for sliced samples and whole tubers using NIR reflectance for Frito Lay 1879 and Russet Norkotah cultivars using 2009 and 2011 combined data.

Sample type	Cultivar _{constituent}	ANN type, characteristics	Training			Testing		
			R _{train}	SeCV _{train} (%)	RPD	R _{test}	Se _{test} (%)	RPD
Slice	FL _{GL}	FFNN, 50	0.58	0.0404	1.21	0.55	0.0370	1.20
	FL _{SU}	FFNN, 50	0.39	0.0455	1.06	0.29	0.0425	1.04
	RN _{GL}	RBFNN	0.84	0.1424	1.65	0.84	0.1596	1.47
	RN _{SU}	NEWGRNN	0.41	0.2700	0.48	0.35	0.2680	0.83
Whole	FL _{GL}	RBFNN	0.68	0.0296	1.13	0.64	0.0301	1.21
	FL _{SU}	RBFNNE	0.77	0.0258	1.45	0.75	0.0258	1.42
	RN _{GL}	RBFNN	0.92	0.0844	2.46	0.82	0.1154	1.57
	RN _{SU}	RBFNNE	0.84	0.0842	1.76	0.83	0.0877	1.64

6.4 Conclusions

NIR reflectance in the range of 900-1685 nm was used to build prediction models using PLSR and different types of artificial neural network for glucose and sucrose sugars in potato tubers that affect quality of French fries and chips. Two cultivars were used to conduct the study, Frito Lay1879 and Russet Norkotah. The study showed promising correlation for both glucose and sucrose using either PLSR or ANN. It should be noted that ANN prediction models were more powerful for sucrose prediction than PLSR, while both methods yielded close results for glucose prediction in the case of Frito Lay 1879 and Russet Norkotah. In general, design radial basis function neural networks (RBFNN) and exact design radial basis function neural networks (RBFNNE) yielded better correlations than feed forward neural networks as the latter type is distinguished for classification and not regression. Sampled wavelengths demonstrated close results to those obtained using full wavelengths and that efficiently reduces the time for data analysis if there is an on-line sorting based on sugars levels. Also, using IPLS and GA as variable selection methods yielded close results to both PLSR and ANN for both cultivars and sugars. However, taking into account that IPLS yielded less variables and yet the same or better performance than GA, and consequently using IPLS saves computation time, and results in a preference of IPLS over GA for variable selection.

Results showed that the classification error obtained from PLSDA models was minimal for FL and RN for glucose more than for sucrose which confirms the prediction results obtained using PLSR as PLSDA is considered the classification tool of PLSR. Whole tubers yielded close classification results compared to sliced samples. In general, Russet Norkotah yielded better correlation than Frito Lay1879 which is possibly due to the fact that RN has higher sugar content than FL as the latter is usually used for processing.

CHAPTER 7 INTEGRATING NIR REFLECTANCE AND VIS/NIR INTERACTANCE SPECTROSCOPIC SYSTEMS DATA (SENSOR FUSION) TO EVALUATE THE PHYSIOLOGICAL STATUS OF POTATO TUBERS

7.1 Introduction

Quality of food products is an important factor by which customers use as a measuring stick to decide which product brand to buy or place from which to get fast food. Chips, French fries, dehydrated, diced and canned potatoes are among the most common products extracted from potatoes.

Near-Infrared (NIR) technology is a rapid, yet accurate technique that has been used to predict quality attributes of agricultural products in sorting, grading, processing, and quality assurance operations of foods. Commercial implementation of NIR spectroscopic systems has been successful in achieving high classification rates for multiple perishable and processed products as shown in section 2.3.3. It was shown in chapters 3-6 that glucose prediction models generally yielded higher correlation statistics than sucrose.

The objective of this study is to investigate the feasibility to integrate data from NIR reflectance and VIS/NIR interactance to predict glucose and sucrose for potato tubers and also classify tubers based on either sugar levels and compare the performance of such fusion with that of the individual modes, i.e. the VIS/NIR interactance and NIR reflectance.

7.2 Materials and Methods

7.2.1 Raw Materials

Two cultivars were chosen to conduct the experiments as discussed in detail in chapter 4.

7.2.2 Data Handling and Analysis

7.2.2.1 Data fusion

Spectroscopic systems are known to be faster in signal acquisition than traditional imaging. Data for this analysis includes relative VIS/NIR interactance data (900-1685 nm) and relative NIR reflectance values (504.8-1004.4 nm). Teflon was used as a reference for the two systems to calculate the relative signals. Interactance and reflectance data were concatenated and each column was then normalized (i.e. each value in a column was divided by the maximum value in the column). It is important to note that only selected wavelengths acquired from the two systems were combined.

7.2.2.2 Data analysis

7.2.2.2.1 Partial least squares regression (PLSR)

A complete description of PLSR used in this research along with pretreatment for either spectra or reference values is listed in section 3.2.4.

7.2.2.2.2 Artificial neural network (ANN)

The ANN types, and configurations applied in this study were the same as that used in section 4.2.4.3.

7.2.2.2.3 Classification of potatoes based on sugar levels

A complete description of the techniques used in potatoes classification based on selected wavelengths is stated in section 5.2.4.5.

7.3 Results and Discussion

7.3.1 Partial Least Squares Regression (PLSR) Results

Results for PLSR for interactance and reflectance combined data from the 2009 and 2011 seasons are shown in table 7.2 with the best prediction results for each season using PLSR shown in table 7.1 which was obtained from chapters 5 and 6 for the interactance and reflectance data sets respectively. For sliced samples, FL best glucose prediction was obtained from the reflectance mode with R(RPD) values of 0.83(1.78) using 2011 season data whereas those values for interactance and reflectance combined data were as close as 0.94(2.84) obtained also from the 2011 season data. However, the best glucose prediction model obtained for RN from reflectance with R(RPD) values of 0.97(4.21) from the 2011 season data did not show an improvement for interactance and reflectance combined data in which R(RPD) values were 0.98(4.97). The best sucrose prediction model for FL was obtained from interactance mode with R(RPD) values of 0.81 (1.70) for 2009 season, while these values for interactance and reflectance combined data were as weaker as 0.62(1.17). Moreover, for RN, the best sucrose prediction model was obtained from reflectance mode with R(RPD) values of 0.81(1.66) from 2011 season data. Such performance was not conducted using the two modes mix data.

In the case of whole tubers, the best glucose prediction model for FL was obtained from interactance data with R(RPD) values of 0.85(1.92) from 2009 season data. Such prediction was weaker using interactance and reflectance data mix with R(RPD) values of 0.67(1.35) from 2011 season data. For RN, R(RPD) values were 0.98(4.84) obtained from reflectance data in the 2011 season. These values were slightly improved using interactance and reflectance mix for the 2011 season with R(RPD) values of 0.98(5.64). For best sucrose prediction model, results for FL showed R(RPD) values of 0.80(1.64) using 2009 interactance data. With the two modes mix data

these values improved to 0.93(2.80) obtained from the 2011 season. For RN, the best sucrose prediction model was obtained from 2011 reflectance mode with R(RPD) values of 0.93(2.77).

Using the two modes data mix for 2011 season data the latter values improved to 0.97(4.23).

Table 7.1. Summary of the best prediction models using PLSR for glucose and sucrose using VIS/NIR interactance and NIR reflectance individual modes for sliced samples and whole tubers for Frito Lay 1879 and Russet Norkotah cultivars.

Sample type	Cultivar _{constituent}	Mode	Validation model		
			R _{pred}	RMSEP	RPD
Slice	FL _{GL}	Reflectance	0.83	0.0435	1.78
	FL _{SU}	Interactance	0.81	0.0391	1.70
	RN _{GL}	Reflectance	0.97	0.0468	4.21
	RN _{SU}	Reflectance	0.81	0.2296	1.66
Whole	FL _{GL}	Interactance	0.85	0.0142	1.92
	FL _{SU}	Interactance	0.80	0.0384	1.64
	RN _{GL}	Reflectance	0.98	0.0387	4.84
	RN _{SU}	Reflectance	0.93	0.1128	2.77

Table 7.2. PLSR results for predicting glucose and sucrose using fused data from VIS/NIR interactance and NIR reflectance systems for sliced samples and whole tubers for Frito Lay 1879 and Russet Norkotah cultivars in the 2009 and 2011 seasons.

Season	Sample type	Cultivar _{constituent}	Preprocessing ^a	Calibration model				Validation model		
				R _{cal}	RMSEC	RMSEC _{cv}	LVs	R _{pred}	RMSEP	RPD
2009	Slice	FL _{GL}	A ₄ ,B ₃ ,C ₁	0.68	0.4874	0.5715	2	0.65	0.5014	1.32
		FL _{SU}	A ₄ ,B ₃ ,C ₂	0.54	0.0676	0.0743	6	0.46	0.0722	1.12
		RN _{GL}	A ₁₂ ,B ₃ ,C ₂	0.99	0.0116	0.0493	2	0.98	0.0194	4.97
		RN _{SU}	A ₀ ,B ₁ ,C ₀	0.65	0.0628	0.0696	20	0.58	0.0664	1.22
	Whole	FL _{GL}	A ₀ ,B ₃ ,C ₀	0.70	0.0186	0.0210	20	0.67	0.0200	1.35
		FL _{SU}	A ₇ ,B ₃ ,C ₂	0.64	0.0554	0.0674	17	0.57	0.0610	1.22
		RN _{GL}	A ₀ ,B ₁ ,C ₀	0.94	0.0325	0.0397	20	0.93	0.0355	2.66
		RN _{SU}	A ₀ ,B ₁ ,C ₀	0.64	0.0631	0.0697	20	0.58	0.0668	1.22
2011	Slice	FL _{GL}	A ₁₂ ,B ₃ ,C ₂	0.98	0.0150	0.0582	8	0.94	0.0272	2.84
		FL _{SU}	A ₀ ,B ₁ ,C ₀	0.83	0.0359	0.0437	20	0.62	0.0497	1.17
		RN _{GL}	A ₇ ,B ₁ ,C ₀	0.97	0.0121	0.0156	20	0.88	0.0291	2.00
		RN _{SU}	A ₀ ,B ₃ ,C ₂	0.92	0.0843	0.1189	20	0.62	0.1978	1.00
	Whole	FL _{GL}	A ₄ ,B ₃ ,C ₀	0.41	0.0504	0.0550	6	0.28	0.0560	1.04
		FL _{SU}	A ₇ ,B ₁ ,C ₀	0.96	0.0154	0.0418	20	0.93	0.0214	2.80
		RN _{GL}	A ₀ ,B ₁ ,C ₀	0.99	0.0257	0.0790	20	0.98	0.0332	5.64
		RN _{SU}	A ₀ ,B ₁ ,C ₁	0.98	0.1651	0.7644	16	0.97	0.2200	4.23

^a See table 3.1 footnote.

7.3.2 Artificial Neural Network (ANN) Results

Results of best prediction models of glucose and sucrose, for FL and RN, obtained from VIS/NIR interactance or NIR reflectance modes using ANN are shown in table 7.3 and obtained from chapter 5 and 6 respectively. Moreover, the results for ANN models resulted from interactance and reflectance data mix for 2009 and 2011 seasons are shown in table 7.4. By comparing equivalent values of R(RPD) in tables 7.3 and 7.4, a general note of a significant decrease of the performance of models for mixed data compared to individual modes with few exceptions in which close results between the two cases was achieved. Glucose prediction models for RN in sliced samples and whole tubers from mixed (fused) data showed R(RPD) values of 0.92(2.246) and 0.98(6.73) obtained from the 2011 season. Such values are close to or better than the values for individual best models. A possible reason for performance decline in ANN using mixed data is the relatively high number of variables (3485) compared to individual mode data.

Table 7.3. Summary of the best prediction models using ANN for glucose and sucrose using VIS/NIR interactance and NIR reflectance individual modes for sliced samples and whole tubers for Frito Lay 1879 and Russet Norkotah cultivars

Sample type	Cultivar _{constituent}	Mode	Testing		
			R _{test}	Sep _{test} (%)	RPD
Slice	FL _{GL}	Reflectance	0.96	0.0086	3.47
	FL _{SU}	Reflectance	0.96	0.0157	3.89
	RN _{GL}	Interactance	0.95	0.0190	3.25
	RN _{SU}	Reflectance	0.97	0.0192	3.92
Whole	FL _{GL}	Reflectance	0.78	0.0190	1.58
	FL _{SU}	Reflectance	0.96	0.0163	3.80
	RN _{GL}	Reflectance	0.96	0.0208	3.72
	RN _{SU}	Reflectance	0.97	0.0830	3.78

Table 7.4. ANN results for predicting glucose and sucrose using fused data from VIS/NIR interactance and NIR reflectance systems for sliced samples and whole tubers for Frito Lay 1879 and Russet Norkotah cultivars in the 2009 and 2011 seasons.

Season	Sample type	Cultivar _{Constituent}	ANN type, characteristics	Training			Testing		
				R_{train}	SeCV _{train} (%)	RPD	R_{test}	Sep _{test} (%)	RPD
2009	Slice	FL _{GL}	FFNN, 1000	0.55	0.02	1.26	0.52	0.0231	1.18
		FL _{SU}	RBFNNE	0.43	0.02	0.99	0.38	0.0229	1.04
		RN _{GL}	FFNN, 300	0.84	0.03	1.74	0.81	0.0303	1.68
		RN _{SU}	RBFNNE	0.67	0.07	1.29	0.52	0.0575	1.18
	Whole	FL _{GL}	RBFNNE	0.75	0.02	1.80	0.67	0.0205	1.35
		FL _{SU}	RBFNNE	0.20	0.02	2.04	0.06	0.0740	1.00
		RN _{GL}	RBFNNE	0.77	0.04	1.32	0.56	0.0403	1.19
		RN _{SU}	RBFNNE	0.18	0.05	1.31	0.19	0.0622	0.98
2011	Slice	FL _{GL}	RBFNNE	0.77	0.02	1.52	0.75	0.0200	1.50
		FL _{SU}	RBFNN	0.56	0.03	0.93	0.42	0.0306	0.96
		RN _{GL}	RBFNNE	0.96	0.03	3.28	0.92	0.0338	2.46
		RN _{SU}	RBFNNE	0.78	0.05	1.48	0.65	0.0636	1.31
	Whole	FL _{GL}	FFNN, 1000	0.15	0.05	1.04	0.06	0.0642	0.99
		FL _{SU}	RBFNNE	0.39	0.05	1.07	0.27	0.0602	1.03
		RN _{GL}	RBFNN	0.98	0.04	4.96	0.98	0.0281	6.73
		RN _{SU}	RBFNNE	0.75	0.04	1.30	0.67	0.0459	1.35

7.3.3 Results for Classification of Potatoes Based on Sugar Levels

Classification for sliced samples and whole tubers was conducted using interactance and reflectance combined data and the results are shown in table 7.6 with the best classification results using individual modes, obtained from chapter 5 and 6 respectively, shown in table 7.5. In the case of sliced samples, classification using fused data based on glucose resulted in an error of 24% for FL and 22% for RN which are higher than the lowest error obtained using individual modes (16%). Sucrose-based classification, however, yielded lower errors than glucose as of 14% for FL and 12% for RN.

Classification results obtained for whole tubers is slightly lower or is similar to the lowest errors obtained from individual modes with error values of 0% for both cultivars based on glucose being 19% and 0% for FL and RN. In the case of sucrose, the error values were 29% for FL and 21% for RN. The above results clarify the advantage of combining data from interactance and reflectance modes and the error values can be enhanced by using broader sugar distribution and higher number of samples, especially in the case of glucose.

Table 7.5. Summary of the best classification results based on glucose and sucrose levels using VIS/NIR interactance and NIR reflectance individual modes for sliced samples and whole tubers for Frito Lay 1879 and Russet Norkotah cultivars.

Sample type	Cultivar _{Constituent}	Mode/ classifier	Testing error (%)
Slice	FL _{GL}	Interactance /PLSDA	16
	FL _{SU}	Interactance / PLSDA	23
	RN _{GL}	Interactance / PLSDA	13
	RN _{SU}	Interactance / PLSDA	18
Whole	FL _{GL}	Interactance / PLSDA	18
	FL _{SU}	Interactance / PLSDA	26
	RN _{GL}	Interactance / PLSDA	0
	RN _{SU}	Reflectance/ LDA	18

Table 7.6. Classification results of sliced samples and whole tubers based on glucose and sucrose levels for Frito Lay 1879 and Russet Norkotah cultivars using multiple classification techniques and VIS/NIR interactance and NIR reflectance combined data sets in the 2009 and 2011 seasons.

Season	Sample type	Cultivar _{Constituent}	Preprocessing ^a	Training error (%)				Testing error (%)				
				LDA	Knn	PLSDA	ANN	LDA	Knn	PLSDA	ANN	Combined classifiers
2009	Slice	FL _{GL}	A ₁₂ ; A ₀ ; A ₇ ; A ₉ ;A ₀	32	32	29	31	35	43	34	36	35
		FL _{SU}	A ₀ ; A ₀ ; A ₀ ; A ₀ ;A ₄	39	37	35	36	43	55	36	43	43
		RN _{GL}	A ₁₂ ; A ₀ ; A ₅ ; A ₁₂ ;A ₁₀	23	23	17	27	23	37	22	23	27
		RN _{SU}	A ₇ ; A ₇ ; A ₁₀ ; A ₇ ;A ₀	35	45	31	14	18	13	12	18	13
	Whole	FL _{GL}	A ₇ ; A ₇ ; A ₁₂ ; A ₄ ;A ₇	28	29	19	31	27	24	20	21	19
		FL _{SU}	A ₀ ; A ₀ ; A ₀ ; A ₅ ;A ₁₂	35	47	30	41	43	39	41	44	38
		RN _{GL}	A ₀ ; A ₅ ; A ₀ ; A ₉ ;A ₇	12	25	20	25	18	19	18	19	19
		RN _{SU}	A ₆ ; A ₀ ; A ₇ ; A ₀ ;A ₇	31	36	31	46	37	31	31	42	31
2011	Slice	FL _{GL}	A ₇ ; A ₇ ; A ₉ ; A ₇ ;A ₁	25	18	7	36	24	26	36	24	31
		FL _{SU}	A ₀ ; A ₀ ; A ₀ ; A ₁₀ ;A ₁₀	30	26	11	29	21	72	14	21	24
		RN _{GL}	A ₀ ; A ₁₀ ; A ₇ ; A ₇ ;A ₄	15	47	13	33	43	30	23	33	33
		RN _{SU}	A ₆ ; A ₀ ; A ₉ ; A ₆ ;A ₀	19	21	10	16	20	23	16	20	23
	Whole	FL _{GL}	A ₁₂ ; A ₀ ; A ₀ ; A ₀ ;A ₀	19	43	3	26	23	34	23	37	23
		FL _{SU}	A ₀ ; A ₀ ; A ₉ ; A ₁₀ ;A ₉	20	46	21	31	29	32	29	39	29
		RN _{GL}	A ₀ ; A ₉ ; A ₄ ; A ₉ ;A ₀	2	26	0	27	27	33	0	27	7
		RN _{SU}	A ₀ ; A ₉ ; A ₇ ; A ₀ ;A ₇	8	45	27	27	21	29	21	43	21

^a See table 3.1 footnote.

7.4 Conclusions

The use of interactance and reflectance combined data for either Frito Lay1879 or Russet Norkotah resulted in an improvement of prediction performance using PLSR of glucose in the case of sliced samples for both cultivars, especially FL. For whole tubers, both cultivars benefited from combining interactance and reflectance data. An improvement in prediction performance for both sugars was achieved especially for sucrose. Results for whole tubers enhance the chances of applying the technique for quality monitoring in industry applications. ANN results for combining modes were not as promising as PLSR which is a consequence from the ability of PLSR to handle collinear data and the factors that need to be adjusted in ANN (i.e. number of neurons in the hidden layer, transfer functions, spread value). The use of combined data in building classification based on sugar levels yielded outstanding results for whole tubers with classification error ranging from 0%-4% for both cultivars and based on glucose and sucrose. Such results followed the prediction models obtained from PLSR especially for RN. Moreover, with the reasonable classification performance achieved for whole tubers, there is a potential for combining the two modes in one system for online sorting of potato tubers based on glucose for RN. However, such a target requires improving the classification rates based on glucose for FL, and based on sucrose for both cultivars. Moreover, more training is needed on different cultivars and various sugar thresholds to obtain a robust, yet accurate sorter that meets industry demands.

CHAPTER 8 OVERALL CONCLUSIONS AND FUTURE WORK

This dissertation research made notable improvement in building prediction and classification models for crucial constituents and physical characteristics of potatoes for growers and processing quality managers. Different spectroscopic systems were used including VIS/NIR interactance (446-1125 nm), NIR transmittance (900-1685 nm), and NIR reflectance (900-1685 nm) as well as VIS/NIR hyperspectral imaging system (400-1000 nm). Experiments were utilized for two cultivars Frito Lay1879 (FL) which is a common chipping cultivar and Russet Norkotah (RN) that is used as a fresh or table cultivar.

Experiments were established over three seasons, 2008 which was aimed to be a preliminary study to investigate the potential of using NIR transmittance, VIS/NIR interactance, and VIS/NIR hyperspectral imaging systems to predict glucose, sucrose, primordium leaf counts, specific gravity and soluble solids using partial least squares regression (PLSR). Another two seasons, 2009 and 2011, only focused on research for measuring glucose and sucrose as those are specifically important in frying process quality. Sampling techniques in the three seasons included 0.5'' (12.7 mm) slices, and whole tubers.

Relative interactance values were calculated for VIS/NIR interactance, and relative reflectance and transmission values were calculated for NIR reflectance and NIR transmittance respectively. Finally, relative values of mean reflectance and curve fitting parameters, extracted from an exponentially decaying curve fitting model, were calculated for the VIS/NIR hyperspectral imaging systems. To extract the most effective wavelengths associated with the prediction of glucose and sucrose for 2009 and 2011 seasons, interval partial least squares (IPLS)

and genetic algorithm (GA) techniques were applied. Calibration and prediction methods were then built using PLSR, and artificial neural network (ANN) that included regular radial basis function neural networks (RBFNN), exact design radial basis function neural networks (RBFNNE), generalized radial basis function neural networks (NEWGRNN), and feed forward neural networks with back propagation (FFNN).

Classification of whole tubers and sliced samples, based on thresholds associated with processing applications, was conducted on the data obtained from the 2009 and 2011 seasons. Classification techniques included linear discriminant analysis (LDA), K-nearest neighbors (Knn), partial least squares discriminant analysis (PLSDA), feed forward artificial neural network, and classifier fusion.

The following main conclusions were deduced from this research:

- 1) A comprehensive study was conducted to review the studies for non-destructive and/or rapid measurements of constituents related to the frying industry, and external and internal quality of fresh tubers. In addition, the most common commercial systems were described and compared from the theory of operations and performance prospective. Application of sorting potato tubers based on constituents levels and/or internal and external defects are feasible with the increasing demand of high quality yet healthy processed foods, and the accelerated developed technology that can maintain fast measurements, durable performance, and high accuracy. A brief view was discussed of the possible future trends in quality evaluation of potato tubers and fried products using noninvasive electronic measurements.
- 2) In the 2008 season, interactance mode demonstrated the best performance for most constituents for FL and RN. PLSR calibration and prediction models showed outstanding

performance in the case of sliced samples for primordium leaf counts with R value of 0.95 for FL and 0.90 for RN and for glucose with R values being as high as 0.90 for FL and 0.95 for RN. Sucrose optimum prediction models had less correlation for both cultivars ($R=0.81$ for FL and 0.63 for RN). Specific gravity showed R values as high as 0.61 for FL and 0.59 for RN. Soluble solids content, however, was the least correlated constituent with maximum values of R of 0.55 for FL and 0.37 for RN. Whole tubers showed general decrease in correlation compared against the sliced samples, especially for RN which brought a conclusion that more studies are required in which broader constituents' distribution exists. In general, results achieved in this study are novel for primordium leaf count that was not achieved before using any spectroscopic system.

- 3) VIS/NIR hyperspectral imaging was used in the 2009 and 2011 seasons to measure glucose and sucrose for sliced samples only as whole tubers yielded low correlation results from preliminary results in the 2008 season. To obtain broad sugar distribution, different soil types, and more storage temperatures were used than in the 2008 season. Glucose and sucrose measurements were conducted by juicing only the tuber tissue penetrated by light in contrast to the 2008 season in which the whole tuber was juiced. Strongly correlated models were obtained for glucose of FL with R values as high as 0.80 and 0.96 for FL and RN. Sucrose prediction however, did not show such high correlation for both cultivars with R values of 0.58 for FL and 0.30 for RN. Selected wavelengths using IPLS and GA showed similar correlation performance compared to the full wavelength models for glucose that yielded to R values up to 0.80 for FL and 0.97 for RN. For sucrose, the R values were as high as 0.54 for FL and 0.38 for RN. Most of best prediction models for both cultivars and for glucose and sucrose were obtained using the

mean reflectance signal and IPLS as the wavelength selection technique. Classification of sliced samples based on glucose or sucrose levels was possible with errors of 19% and 18% for FL and RN using glucose thresholds. Classification errors based on sucrose models (34% and 38 for FL and RN) were higher than errors obtained for glucose which followed the same results achieved by PLSR. Prediction and classification results can be improved using broader sugar distribution, using other classifiers such as artificial neural network, and majority voting classification techniques. Selected wavelengths used to build classification results for glucose could be further applied in a multispectral sorting system that may be combined with a computer vision system to obtain multi-tasking sorting for defect detection and also sugar-based sorting.

- 4) VIS/NIR interactance was used in the 2009 and 2011 seasons to study the rapid and/or non-destructive determination of glucose and sucrose for potato tubers for Frito Lay 1879 and Russet Norkotah. Non-noisy wavelengths (2107) and sampled (386) were used to build prediction models using PLSR and ANN. Also IPLS and GA were applied to extract the wavelengths related to best prediction models for both sugars. Encouraging correlation was achieved for FL and RN for both sugars with R values for sliced samples being as high as 0.92 and 0.94 for FL and RN in the case of glucose and 0.82, and 53 for FL and RN in the case of sucrose. Whole tubers prediction models also yielded R values of 0.85 and 0.97 for FL and RN for glucose, and 0.46 and 0.63 for FL and RN for sucrose. General improvement of correlation for sucrose was obtained using selected variable models with R values reaching 0.81 and 0.78 for FL and RN for sliced samples, and 0.80 and 0.94 for RN for whole tubers. Glucose prediction models based on selected variables showed similar performance compared to full wavelength models. The IPLS

method resulted in less number of wavelengths (11-68) for sliced samples, and 20-75 for whole tubers. Whereas GA resulted in significantly higher number of wavelengths of 165-247 for sliced samples, and 182-229 for whole tubers. Consequently, using IPLS in wavelengths selection is much more efficient and less time consuming than GA if they both produce the same performance. In general, a reduction of the number of wavelengths to 0.5-3% of the full wavelengths (2107) was achieved using either IPLS or GA. Classification of whole tubers based on glucose levels yielded errors of 18% and 0% for FL and RN. Whereas the values were 16% and 13% for FL and RN for sliced samples. Classification based on sucrose was weaker, for FL, and classification errors of 26% and 14% for FL and RN were achieved for whole tubers which were similar to the values obtained for sliced samples (23% and 18% for FL and RN).

- 5) NIR diffuse reflectance was utilized on both sliced samples and whole tubers for both the 2009 and 2011 seasons. PLSR and ANN were applied on full (784), sampled (262), and selected wavelengths using IPLS and GA. Prediction results were promising for both sugars in which the R values for sliced samples reached as high as 0.96 and 0.97 for FL and RN for glucose, and for sucrose the values were 0.95 for FL and 0.97 for RN. For whole tubers, R values for glucose prediction models were as high as 0.76 and 0.98 for FL and RN. Moreover, sucrose prediction models also showed high correlation with R values of 0.96 and 0.97 for FL and RN. Prediction results based on sampled wavelengths showed similar performance in most cases compared to full wavelengths models. In addition, using IPLS and GA, similar or better correlation performance, compared to the full wavelength models, was achieved for both sugars and cultivars which indeed clarifies the strength and efficiency of the selected wavelengths in holding sufficient information

about glucose and sucrose. Tubers and sliced samples classification based on sugar levels was not as powerful as prediction models. Whole tubers showed classification error of 19% and 0% for FL and RN based on glucose, and 31% and 21% based on sucrose. Sliced samples showed classification error of 17% and 17% based on glucose and 25% and 18% based on sucrose. Some of the classification results still do not meet industry requirements and performance enhancement can likely be achieved by increasing the number of samples, obtaining broader sugar distribution, and using other classification techniques such as support vector machines (SVM).

- 6) Data combined/fused from VIS/NIR interactance and NIR reflectance resulted in improvement in the case of sliced samples for the prediction of glucose for FL and RN with R values reaching 0.94 and 0.98 respectively. Whole tubers also showed significant improvement in the performance of sucrose prediction with R values as high as 0.93 and 0.97 for FL and RN. Similar performance of classification results, compared to individual modes, was obtained for whole tuber glucose-based models. However, significant improvement was achieved in the case of sucrose-based models for sliced samples with error values of 14% and 12% for FL and RN. Such results indicate that combining data from both modes can lead to more valuable information to explain the variation between samples and enhance classification as well as prediction performance.

This research in general resulted in a promising prediction performance of glucose, sucrose, and primordium leaf counts using different regression techniques, and it represents a basic study that indeed is comparable in performance to previous studies conducted to measure sugars in potato tubers (Dull et al., 1989; Mehrubeoglu and Cote, 1997; Hartman and Buning-Pfaue, 1998; Scanlon et al., 1999; Yaptenco et al., 2000; Haase, 2004; Chen et al., 2005; Subedi

and Walsh, 2009). Moreover, classification of sliced samples based on sucrose showed lower error than the results obtained from individual modes.

However, to obtain more reliable results that tightly relate this study to the practical field, several recommendations can be drawn for future research as follows:

- 1) This study was designed to be a foundation for establishing a handheld device that works with either sliced samples or whole tubers. Another long-term target was also contributing toward the design of on-line sorting systems for potato tubers. Testing tubers under actual field conditions, however, is more difficult than testing in the lab as there is a possibility for having factors such as clay particles and/or moisture on tuber surface that reduce, or even suppress, signal acquired from tubers. If the sorting system is used after harvesting to eliminate or separate tubers with undesired levels of sugars for further reconditioning, there is a need to tackle the presence of clods, rocks, and vine parts in the flow of tubers either by adding a computer vision system to eliminate foreign materials and then following such by the constituent-based sorting mechanism.
- 2) An on-line system or sorting requires working with movable objects, and under such circumstances, a possible reduction of signal quality (signal to noise ratio) acquired from tubers is likely to occur, and consequently lower the performance of constituent prediction-based sorting. Thus, proper choice of the optical components should be taken into account to obtain commercially-accepted functionality and productivity.

Prediction models for glucose, sucrose, and primordium counts obtained from different systems in this study, especially VIS/NIR interactance, and NIR reflectance, or merged data between the interactance and reflectance modes were encouraging. However, the change in

spatial distribution of some constituents over the storage period obligates conducting more experiments with different cultivars, growing and storage conditions to confirm the obtained prediction and classification models so that more robust, reproducible, and stable performance can be later applied on commercial systems.

BIBLIOGRAPHY

BIBLIOGRAPHY

- Abbott, J.A., (1999). Quality measurement of fruits and vegetables. *Postharvest Biology and Technology*, 15: 207-225.
- Agle, W.M., Woodbury, G.W. (1968). Specific gravity-dry matter relationship and reducing sugar changes affected by potato variety, production area, and storage. *American Journal of Potato Research*. 45(4): 119-131.
- Allen, E.J., O'Brien, P.J., Firman, D. (1992). Seed tuber production and management. In P. Harris (Ed.), *the potato crop* (pp. 247-291). Chapman & Hall, London, UK.
- Al-Mallahi, A., Kataoka, T. Okamoto, H. Shibata, Y. (2010a). An image processing algorithm for detecting in-line potato tubers without singulation. *Computers and Electronics in Agriculture*, 70(1): 239–244.
- Al-Mallahi, A., Kataoka, T. Okamoto, H. Shibata, Y. (2010b). Detection of potato tubers using an ultraviolet imaging-based machine vision system. *Biosystems Engineering*, 105(2): 257–265.
- Al-Mallahi, A., Kataoka, T. Okamoto, H. (2008a). Discrimination between potato tubers and clods by detecting the significant wavebands. *Biosystems Engineering*, 100(3): 329–337.
- Al-Mallahi, A., Kataoka, T., Okamoto, H. (2008b). An algorithm for distinguishing potato tubers on the conveyor of the potato harvester using uv camera. *Proceedings the 2008 ASABE annual international meeting*, Rhode Island, USA.
- Balls, R.C., Gunn, J.S., Starling, A.J. (1982). *The national potato damage awareness campaign. Potato marketing broad and agricultural development and advisory service*. Oxford, 32pp.
- Baritelle, A.L., Hyde, G.M. (1999). Effect of tuber size on failure properties of potato tissue. *Transactions of the ASAE*, 42 (1): 159–161.
- Baritelle, A.L., Hyde, G.M., Thornton, R.E., Bajema, R.W. (1998). A classification system for impact related defects in potato tubers. *Proceedings of the 38th annual Washington State Potato Conference and Trade Fair*, Feb., 1999.
- Barnes, M., Duckett, T., Cielniak, G. (2009). Boosting minimalist classifiers for blemish detection in potatoes. 24th International Conference Image and Vision Computing, New Zealand (IVCNZ 2009).
- Barnes, M., Duckett, T., Cielniak, G., Stroud, G., Harper, G. (2010). Visual detection of blemishes in potatoes using minimalist boosted classifiers. *Journal of Food Engineering*, 98(3): 339–346.

- Barreiro, P., Ruiz-Altisent, M., Fdez-Valle, E., Ruiz-Cabello, J., Recasens, I., Asensio, M. (1998). Mealiness assessment in apples and peaches using mri techniques. In: AgEng98, International Conference on Agricultural Engineering, Oslo. CIGR, Brussels paper 98-F-074.
- Barton, F.E., Kays, S.E. (2001). Analytical application to fibrous foods and commodities. In P. Williams, K. Norris (Eds.), near-Infrared technology in the agricultural and food industries (pp. 215-231). American Association of Cereal Chemists, USA.
- BeMiller, J.N. (2003). Carbohydrate analysis. In N.N. Nielsen (Ed.), food analysis (pp. 143-174). Kluwer Academic/ Plenum Publishers, USA.
- BeMiller, J.N. (2010). Carbohydrate analysis. In S. S. Nielsen (Ed.), food analysis (pp. 147-177). Springer, NY, USA.
- Ben-Gera, I., Norris, K.H. (1968). Direct spectrophotometric determination of fat and moisture in meat products. *Journal of Food Science*, 33(1): 64-67.
- Birth, G.S. (1960). A nondestructive technique for detecting internal discolorations in potatoes. *American Potato Journal*, 37(2): 53-60.
- Bishop, C.M. (2007). Pattern recognition and machine learning. Springer Science+ Business Media, LLC, NY, USA.
- Bohl, W.H., Johnson, S.B. (2010). Commercial potato production in North America. The potato association of America handbook, second revision of American Potato Journal supplement volume 57 and USDA handbook 267 by the extension section of the Potato Association of America.
- Brecht, J.K., Shewfelt, R.L., Garner, J.C., Tollner, E.W. (1991). Using x-ray computed tomography (x-ray CT) to nondestructively determine maturity of green tomatoes. *HortScience*, 26: 45–47.
- Brook, R.C. (1996). Potato bruising; how and why; emphasizing blackspot bruise. Running Water Publishing, Haslett, MI, USA.
- Brown, P., Russell, C. (2001). Irrigation and tuber susceptibility to bruising. Tasmanian Institute of Agricultural Research. Australia.
- Brunt, K., Drost W.C. (2010). Design, construction, and testing of an automated nir in-line analysis system for potatoes. Part I: off-line nir feasibility study for the characterization of potato composition. *Potato Research*, 53(1): 25–39.
- Burton, W.G. (1989). The potato (3rd Ed., pp. 84-155). Longman scientific and technical, London, UK.

- Burton, W.G., Van Es, A., Hartmans, K.J. (1992). The physics and physiology of potato storage. In By P. Harris (Ed.), the potato crop the scientific basis for improvement (2nd Edn., pp. 428-475). Chapman & Hall, London, UK.
- Butz, P., Hofmann, C., Tauscher, B. (2005). Recent developments in noninvasive techniques for fresh fruit and vegetable internal quality analysis. *Journal of Food Science*, 70 (9): 131-141.
- Cen, H. (2011). Hyperspectral imaging-based spatially-resolved technique for accurate measurement of the optical properties of horticultural products. Dissertation, Michigan State University, East Lansing, MI, USA.
- Chang, C-I. (2007). Overview. In Chein-I Chang (Ed.), hyperspectral data exploitation theory and applications (pp. 1-16). Willey-Interscience. USA.
- Chang, S.K.C. (2010). Protein analysis. In S. Suzanne Nielsen (Ed.), food analysis (4th Edn., pp. 133-146), Springer, NY, USA.
- Chao, K. (2010). Automated poultry carcass inspection by a hyperspectral– multispectral line-scan imaging system. In D-W Sun (Ed.), hyperspectral imaging for food quality analysis and control (pp. 241-271). Elsevier Press, UK.
- Chen J.Y., Zhang H., Miao Y., Matsunaga R. (2005). NIR measurement of specific gravity of potato. *Food Science Technology Research*, 11(1):26-31.
- Chen, J.Y., Miao, Y., Zhang, H., Matsunaga, R. (2004). Non-destructive determination of carbohydrate content in potatoes using near infrared spectroscopy. *Journal of Near Infrared Spectroscopy*, 15(5): 311-314.
- Chen, J.Y., Zhang H., Miao Y., Asakura, M. (2010). Nondestructive determination of sugar content in potato tubers using visible and near infrared spectroscopy. *Japan Journal of Food Engineering*, 11(1): 59-64.
- Chen, P. (1978). Use of optical properties of food materials in quality evaluation and materials sorting. *Journal of Food Process Engineering*, 2(4): 307-322.
- Chen, P., Sun, Z. (1991). A review of non-destructive methods for quality evaluation and sorting of agricultural products. *Journal of Agricultural Engineering Research*, 49: 85-98.
- Chen, Y-R., Chao, K., Kim, M.S. (2002). Machine vision technology for agricultural applications. *Computers and Electronics in Agriculture*, 36: 173-191.
- Cheng, Y., Haugh, C.G. (1994). Detecting hollow heart in potatoes using ultrasound. *Transactions of the ASAE*, 37(1): 217–222.

- Chenglong, W., Xiaoyu, L., Wei, W., Jie, L., Hailong, T., Dongdong, W. (2011). Detection of potato's size based on centroidal principal axis. *African Journal of Agricultural Research*, 6(17): 4140–4148.
- Christy, A. A., Kvalhiem, O. M. (2007). Latent-variable analysis of multivariate data in infrared spectrometry. In Y. Ozaki, W.F. McClure, A. A. Christy (Eds.), *near-infrared spectroscopy in food science and technology* (pp.145-162). Wiley-Interscience, USA.
- Coles, G.D., Lammerink, J.P., Wallace, A.R. (1993). Estimating potato crisp colour variability using image analysis and a quick visual method. *Potato Research*, 36: 127- 134
- Conway, J.M. Norris, K.H. Bodwell, C.E. (1984). A new approach for the estimation of body composition: infrared interactance. *American Journal of Clinical Nutrition*, (40): 1123-1130.
- Cutter, E. (1992). Structure and development of the potato plant. In P. Harris (Ed.), *the potato crop* (pp 65-161). Chapman and Hall, U.K.
- Dacal-Nieto, A., Formella, A., Carrión, P., Vazquez-Fernandez, E., Fernández-Delgado, M. (2011). Common scab detection on potatoes using an infrared hyperspectral imaging system. *Proceedings of the 16th International Conference on Image Analysis and Processing - volume part ii*, 303–312.
- Dacal-Nieto, A., Vazquez-Fernandez, E., Formella, A., Martin, F., Torres-Guijarro, S., Gonzalez-Jorge, H. (2009). A genetic algorithm approach for feature selection in potatoes classification by computer vision. *Proceedings the 35th Annual Conference of IEEE Industrial Electronics*, Porto, Portugal, 1955–1960.
- Dahm, D.J., Dahm, K.D. (2001). The physics of near-infrared scattering. In P. Williams and K. Norris (Eds.), *near-infrared technology in the agricultural and food industries* (pp.1-18). ASCC., St. Paul, Minnesota, USA.
- Dean, B. (1996). The chemical nature of blackspot bruising. In R. C. Brook, (Ed.), *potato bruising; how and why; emphasizing Blackspot bruise*. Running Water publishing, Haslett, MI, USA.
- Deck, S.H., Morrow, C.T., Heinemann, P.H., Sommer III, H.J. (1995). Comparison of a neural network and traditional classifier for machine vision inspection of potatoes. *Applied Engineering in Agriculture*, 11(2): 319–326.
- Donis-González, I.R., Guyer, D.E., Pease, A., Barthel, F. (2014). Internal characterisation of fresh agricultural products using traditional and ultrafast electron beam X-ray computed tomography imaging. *Biosystems Engineering*, 117: 104-113.
- Donis-Gonzalez, I.R., Guyer, D.E., Pease, A., Fulbright, D.W. (2012). Relation of computerized tomography Hounsfield unit measurements and internal components of fresh chestnuts (*Castanea spp.*). *Postharvest Biology and Technology*, 64(1): 74-82.

- Dufour, E. (2009). Fundamentals and instruments. In D-W Sun (Ed.), *infrared spectroscopy for food quality analysis and control*. Academic press, Elsevier Press, USA.
- Dull, G.G., Birth, G.S., Leffler, R.G. (1989). Use of near infrared analysis for the nondestructive measurement of dry matter in potatoes. *American Potato Journal*, 66: 215-225.
- Ebrahimi, E., Mollazade, K., Arefi, A. (2011). Detection of greening in potatoes using image processing techniques. *Journal of American science*, 7(3): 243–247.
- Economic Research Services (ERS) (2012). Vegetables pulses: potatoes. Available at: <http://www.ers.usda.gov/topics/crops/vegetables/pulses/potatoes.aspx#.UtQ9fm086So>.
- El batawi, I.E. (2008). An acoustic impact method to detect hollow heart of potato tubers. *Biosystems Engineering*, 100(2): 206–213.
- El Masry, G., Cubero, S., Moltó, E., Blasco, J. (2012). In-line sorting of irregular potatoes by using automated computer-based machine vision system. *Journal of Food Engineering*, 112 (1–2): 60–68.
- El Masry, G., Sun, D.-W. (2010b). Meat quality assessment using a hyperspectral imaging system. In D.-W. Sun (Ed.), *hyperspectral Imaging for food quality analysis and control* (pp.175-240). Elsevier Press, UK.
- El Masry, G., Sun, D-W. (2010a). Principles of hyperspectral imaging technology. In D-W Sun (Ed.), *hyperspectral imaging for food quality analysis and control* (pp.3-43). Elsevier Press, UK.
- Esehaghbeygi, A., Raghmi, N., Kargar, A. (2011). Detection of internal defects in potato based on ultrasound attenuation. *American Journal of Potato Research*, 88(2): 160–166. .
- Evans, S.D. (1999). Reflectance spectrophotometry of bruising in potatoes. II. Acquisition of spectra in the visible spectrum with an integrating sphere. *Int. Agrophysics*, 13: 203-209.
- FAOSTAT. (2013). Potatoes Production in the World. Statistics Division. Available at: www.faostat3.fao.org/home/index.html#visualize_by_domain. accessed, September 2012.
- Faulkner, G. (2012). Essential trends in world potato markets. *World Potato Markets*, Agricultural Markets Ltd, UK.
- Fernandez-Ahumada, E., Roger, J.M., Palagos, B., Guerrero, J.E., Perez-Marín, D., Garrido-Varo, A. (2010). Multivariate near-infrared reflection spectroscopy strategies for ensuring correct labeling at feed bagging in the animal feed industry. *Applied Spectroscopy*, 64 (1): 83–91.
- Finney, E.E., Norris, K.H. (1973). X-ray images of hollow heart potatoes in water. *American Potato Journal*, 50(1): 1–8.

- Finney, E.E., Norris, K.H. (1978). X-ray scans for detecting hollow heart in potatoes. *American Potato Journal*, 55(2): 95–105.
- Fraser, D.G., Jordan, R.B., Künemeyer, R., McGlone, V.A. (2003). Light distribution inside mandarin fruit during internal quality assessment by nir spectroscopy. *Postharvest Biology and Technology*, 27: 185–196.
- Fraser, D.G., Künemeyer, R., McGlone, V.A., Jordan, R.B., 2000. Near infrared light penetration into an apple. *Postharvest Biology and Technology*, 22: 191–194.
- Friedman, Ed., Miller, J.L. (2003). Introduction. In E. Friedman, J.L. Miller (Eds.), *photonics rule of thumb optics, electro-optics, fiber optics, and laser* (pp. xi-xiv). SPIE Press, McGraw-Hill. WA, USA.
- Gaze, S.R., Stalham, R.M., Newbery, R.M., Allen, E.J. (1998). BPC project report 807/182 (1997). British potato council. Oxford, U.K.
- Giambattista, A., Richardson, M.M., Richardson, R.C. (2007). *College physics* (2nd edn., pp. 802-836). Mc Graw Hill Higher Education, NY, USA.
- Glynne, M.D., Jackson, V.G. (1919). The distribution of dry matter and nitrogen in the potato tubers. Variety King Edward. *Journal of Agricultural Science*, 9(3): 237-258.
- Gottschalk, K., and Ezekiel, R. (2006). Storage. In J. Gopal, S.M.P. Khurana (Eds.), *handbook of potato production, improvement, and postharvest management* (pp. 489-522). The Haworth Press Inc., NY, USA.
- Gould, W.A. (1995a). Specific gravity-its measurement and use. In W.A. Gould, J.R. Sowokinos, E. Bantari, P.H. Orr, D.A. Preston (Eds.), *chipping potato handbook*. (pp. 18-21). The Snack Food Association, VA, USA.
- Gould, W.A. (1995b). Color and color measurement of potato chips. In W.A. Gould, J.R. Sowokinos, E. Bantari, P.H. Orr, D.A. Preston (Eds.), *chipping potato handbook*. (pp. 30-31). The Snack Food Association, VA, USA.
- Gray, D, Hughes, J.C. (1978). Tuber quality. In . P.M. Harris (Ed.), *the potato crop* (pp.504–544). Chapman and Hall, London, UK.
- Grunenfelder, L., Hiller, L.K., Knowles, N.R. (2006). Color indices for the assessment of chlorophyll development and greening of fresh market potatoes. *Postharvest Biology and Technology*, 40(1): 73–81.
- Ha, K-L., Kanai, H., Chubachi, N., Kamimura, K. (1991). A Basic study on nondestructive evaluation of potatoes using ultrasound. *Japanese Journal of Applied Physics* 30(30–1): 80–82.

- Haase, N.U. (2004). Estimation of dry matter and starch concentration in potatoes by determination of under-water weight and near infrared spectroscopy. *Potato Research*, 46(3-4): 117-127.
- Haase, N.U. (2006). Rapid estimation of potato tuber quality by near-infra red spectroscopy. *Starch Journal*. 58(6): 268-273.
- Haase, N.U. (2011). Prediction of potato processing quality by near infrared reflectance spectroscopy of ground raw tubers. *Journal of Near Infrared Spectroscopy*, 19(1): 37-45.
- Hartmann, R., Binng-Pfaue, H. (1998). Nir determination of potato constituents. *Potato Research*, 41(4): 327- 334.
- Harvey, R.B. (1937). The x-ray inspection of internal defects of fruit and vegetables. *American Potato Journal*, 35: 156-157.
- Hasankhani, R., Navid, H. (2012). Potato sorting based on size and color in machine vision system. *Journal of Agricultural Science*, 4(5): 235-244.
- Hasankhani, R., Navid, H., Seyedarabi, H. (2012). Potato surface defect detection in machine vision system. *African Journal of Agricultural Research*, 7(5): 844–850.
- Haverkort, A.J., Van Loon, C.D., Van Eijck, P., Scheer, F.P., Schijvens, E.P.H.M., Uitslag, H., Baarveld, H.R., Campobello, E.W.A., Liefrink, S.R. Peeten, H.M.G. (2002). *On the road to potato processing* (2nd Edn.), The Netherlands Consultative Potato Institute, Netherlands.
- Haykin, S. (2009). *Neural network and learning machines* (3rd Edn.). Pearson Education, Inc., NJ, USA.
- Heinemann, P.H., Pathare, N.P., Morrow, C.H. (1996). An automated inspection station for machine-vision grading of potatoes. *Machine Vision and Applications*, 9(1): 14–19.
- Helgerud T., Segtnan, V.H., Wold, J.P., Balance, S., Knutsen, S.H., Rukke, E.O., Afseth, N.K. (2012). Near-infrared spectroscopy for rapid estimation of dry matter content in whole unpeeled potato tubers. *Journal of Food Research*, 4: 55-65.
- Hemmat, A., Taki, O. (2001). Potato losses and mechanical damage by potato diggers in the fereidan's region of Isfahan. *J. Science and Technology Agriculture and Natural Resources*, 5(2): 195–209 (in Farsi with English Abstract).
- Hernández-Sánchez, N., Hills, B., Barreiro, P., Marigheto, N. (2007). An nmr study on internal browning in pears. *Postharvest Biology and Technology*, 44(3): 260-270.
- Hoffmann, T., Wormans, G., Fürll, C. (2005). A system for determining starch in potatoes online. *IAG Eng LUA and LU of Ag.*, 37(2): 34-43.

- Hopkins, R.B. (1953). The reduction of injuries to potato tubers through the use of padding materials. *American Potato Journal*, 30(10): 247–255.
- Hosainpour, A., Komarizade, M.H., Mahmoudi, A., Shayesteh, M.G. (2010). Feasibility of impact-acoustic emissions for discriminating between potato tubers and clods. *Journal of Food, Agriculture, and Environment*, 8(2):565-569.
- Houghland, G.V.C. (1966). New conversion table for specific gravity, dry matter, and starch in potatoes. *American Potato Journal*, 43: 138.
- Hughes, J.C. (1980). Role of tuber properties in determining susceptibility of potatoes to damage. *Annals of Applied Biology*, 96: 344–345.
- Hyde, G.M., Thornton, R.E., Iritani, W.M. (1979). Chain speed effects on potato tuber damage and soil elimination. *ASAE paper No. 79 – 3014*.
- Jeong, Jin-C., Ok, Hyun-C., Hur, On-S., Kim, Chung-G. (2008). Prediction of sprouting capacity using near-infrared spectroscopy in potato tubers. *American Journal of Potato Research*, 85(5): 309-314.
- Jin, J., Li, J., Liao, G., Yu, X., Viray, L.C.C. (2009). Methodology for potatoes defects detection with computer vision. *Proceedings of the 2009 International Symposium on Information Processing*, 346–351.
- Jivanuwong, S. (1998). Nondestructive detection of hollow heart in potatoes using ultrasonic. A master thesis. Department of Biological Systems Engineering. Virginia Polytechnic Institute and State University, VA, USA.
- Jobling, J. (2000). Potatoes: handle with care. Sydney postharvest laboratory information sheet. *Good Fruit and Vegetables*, 11(4): 34–35.
- Johnston, F.B, Hoffman, I., Petrasovits, A. (1968). Distribution of mineral constituents and dry matter in the potato tuber. *American Potato Journal*, 45: 287–292.
- Kadam, S.S., Dhumal, S.S., Jambhale, N.D. (1991a). Structure, nutritional composition, and quality. In D.K. Salunkhe, S.S. Kadam, and S.J. Jadhav (Eds.), *potato: production, processing, and products*, (pp.9-36), CRC press, USA.
- Kadam, S.S., Wankier, B.N., Adsule, R.N. (1991b). Processing. In D.K. Salunkhe, S.S. Kadam, S.J. Jadhav (Eds.), *potatoes: productions, processing and products* (pp. 111-154). CRC Press, USA.
- Kang S., Lee, K-J., Choi, W., Son, J-R., Choi, D-S., Kim, G. (2003). A near-infrared sensing technique for measuring the quality of potatoes. *Proceedings the ASAE Annual International Meeting*. Las Vegas, Nevada, USA.

- Kang, S., Lee, K.-J., Son, J.-R. (2008). Online internal quality evaluation system for the processing potatoes. ASABE Publication No. 701P0508cd. Proceedings of the ASABE Food Processing Automation Conference, Providence, Rhode Island, USA.
- Kärenlampi, S.O., White, P.J. (2009). Potato proteins, lipids, and minerals. In J. Singh, L. Kaur (Eds.), *advances in potato chemistry and technology* (pp. 99-126), Academic Press, Elsevier, NY, USA.
- Kawano S. (2002). Applications to agricultural products and foodstuffs. In H.W. Siesler, Y. Ozaki, S. Kawata, H.M. Heise (Eds.), *near-infrared spectroscopy principles, instruments, applications* (pp. 115–124). Weinheim : Wiley-VCH, USA.
- Kellock, T. (1995). Potatoes: factors affecting dry matter. Department of Primary Industries. Note No. AG0323. Melbourne, Australia.
- Kerekes, J.P., Schott, J.R. (2007). Hyperspectral imaging systems. In C.-I. Chang (Ed.), *hyperspectral data exploitation: theory and applications* (pp. 19-45). Wiley-Interscience, USA.
- Kirk, W.W., Davies, H.V., Marshall, B. (1985). The effect of temperature on the initiation of leaf primordia in developing potato sprouts. *Journal of Experimental Botany*, 36(171): 1634-1643.
- Klapp, E. (1945). *Kartoffelbau: wesen und praktischer anbau der kartoffel einschliesslich des treibkartoffel- und pflanzgutbaues* Stuttgart.
- Kleinhenz, M. (2001). The tuber times, potato growing tips and news from the world of research. Ohio State University extension. Oh, USA.
- Kleinschmidt, G.D., Thornton, M.K. (1991). Bruise free potatoes: Our goal, Bulletin 725, university of Idaho cooperative extension system, Moscow, Idaho, USA.
- Kunkel, R, Gardner, W.H. (1965). Potato tuber hydration and its effect on black spot of Russet Burbank in the Columbia Basin of Washington. *American potato Journal*, 42: 109 –124.
- Lammertyn, J., Peirs, A., De Baerdemaeker, J., Nicolai, B.M. (2000). Light penetration properties of nir radiation in fruit with respect to non-destructive quality assessment. *Postharvest Biology and Technology*, 18: 121–132.
- Lamp, K. (1960). Die widerstandsfähigkeit von kartoffelknollen gegen beschädigung. *European Potato Journal*, 3: 13–29.
- Lawrence, K.C., Windham, W.R., Park, B., Buhr, R.J. (2001). Hyperspectral imaging for poultry contaminant detection. *Near Infrared Spectroscopy News*, 12(5): 3-6.

- Liu, Q., Donner, E., Tarn, R., Singh, J., Chung, H-J. (2009). Advanced analytical techniques to evaluate the quality of potato and potato starch. In J. Singh, I kaur (Eds.), advances in potato chemistry and biotechnology (pp. 221-248). Academic Press, Elsevier, USA.
- Loow, H. (1964). Mechanical damages to potatoes. Bulletin No. 304, Swedish Institute of Agricultural Engineering, Uppsala. Sweden.
- Lu, R., Peng, Y. (2006). Hyperspectral scattering for assessing peach fruit firmness. *Biosystems Engineering*, 93(2):161-171.
- Mahendran R., Jayashree G.C., Alagusundaram K. (2012). Application of computer vision technique on sorting and grading of fruits and vegetables. *Journal of Food Process and Technology*. S1-001. doi:10.4172/2157-7110.S1-001.
- Marchant, J.A., Onyango, C.M., Street, M.J. (1990). Computer vision for potato inspection without singulation. *Computers and Electronics in Agriculture*, 4(3): 235–244.
- Marique, T., Kharoubi, A., Bauffe, P., Ducattillon, C. (2003). Modeling of fried potato chips color classification using image analysis and artificial neural network. *Food Engineering and Physical Properties*, 48(7): 2263-2266.
- Marique, T., Pennincx, S., Kharoubi. A. (2005). Image segmentation and bruise identification on potatoes using a kohonen self-organizing map. *Journal of Food Science and Physical Properties*, 70(7): 415–417.
- Mathanker, S.K., Weckler, P.R., Bowser T.J. (2013). X-ray applications in food and agriculture: a review. *Transactions of the ASABE*, 56(3): 1227-1239.
- Mathew, R., Hyde G. M. (1997). Potato impact damage thresholds. *Transactions of the ASAE*, 40(3): 705–709.
- McCay, C.M., McCay, J.B., Smith, O. (1975). The nutritive value of potatoes. In W.F. Talburt O. Smith (Eds.), potato processing (pp. 235-274). The AVI Publishing Company, Inc., USA.
- McClure, W.F. (2007). Introduction. In Y. Ozaki, W.F. McClure, A.A. Christy (Eds.), near-infrared spectroscopy in food science and technology (pp.1-10). Wiley-Interscience, USA.
- Mehl, P.M., Chao, K., Kim, M., Chen, Y.R. (2002). Detection of defects on selected apple cultivars using hyperspectral and multispectral image analysis. *Applied Engineering in Agriculture*, 18(2): 219–226.
- Mehrotra, K., Mohan, C., Ranka, S. (1996). Elements of artificial neural networks. A Bradford Book, MIT Press, USA.

- Mehrubeoglu, M., Cote, G.L. (1997). Determination of total reducing sugars in potato samples using near-infrared spectroscopy. *Cereal Foods World*, 42(5): 409-413.
- Mendoza, F., Dejmek, P., Aguilera, J.M. (2007). Colour and image texture analysis in classification of commercial potato chips. *Food Research International*, 40: 1146–1154.
- Menesatti, P., Costa, C., Aguzzi, J. (2010). Quality evaluation of fish by hyperspectral imaging. In D-W Sun (Ed.), *hyperspectral imaging for food quality analysis and control* (pp.273-294). Elsevier Press, UK.
- Milczarek, R., Saltveit, M., Garvey, T., McCarthy, M. (2009). Assessment of tomato pericarp mechanical damage using multivariate analysis of magnetic resonance images. *Postharvest Biology and Technology*, 52(2): 189-195.
- Miller, W.M., Stephenson, K.Q. (1971). Vibrational response of potatoes as a sorting criterion. ASAE Paper No. 71-652.
- Mizrach, A. (2012). Ultrasound for fruit and vegetable quality evaluation. In: *handbook on applications of ultrasound sonochemistry for sustainability* (pp. 129-162). CRC Press. Taylor and Francis Group, FL, USA.
- Mizrach, A. (2008). Ultrasonic technology for quality evaluation of fresh fruit and vegetables in pre- and postharvest processes. *Postharvest Biology and Technology*, 48: 315–330.
- Mizrach, A., Galili, N., Rosenhouse, G. (1992). Half-cut fruit response to ultrasonic excitation. ASAE Paper No. 923017. American Society of Agricultural Engineers, St. Joseph, MI.
- Mizrach, A., Galili, N., Rosenhouse, G. (1989). Determination of fruit and vegetable properties by ultrasonic excitation. *Transactions of the ASAE*, 32: 2053–2058.
- Mohr, W.P., Spurr, A.R., Fenn, P., Timm, H. (1984). X-ray microanalysis of hollow heart potatoes. *Food Microstructure*, 3(1): 41–48.
- Mohsenin, N.N. (1986). *Physical properties of plant and animal materials*. Gordon and Breach science publishes. New York, USA.
- Molto, E., Blasco, J., Gomez-Sanchiz, J. (2010). Analysis of hyperspectral images of citrus fruits, In D.-W. Sun (Ed.), *hyperspectral imaging for food quality analysis and control* (pp.321-348). Elsevier Press, UK.
- Mottram, D.S., Wedzicha, B.L. (2002). Acrylamide is formed in the Maillard reaction. *Nature (International Weekly Journal of Science)*, 419: 448-449.
- Muir, A.Y., Porteos, R.L., Wastie, R.L. (1982). Experiments in the detection of incipient diseases in potato tubers by optical methods. *Journal of Agricultural Engineering Research*, 2: 131-138.

- Muir, A.Y., Ross, D.W., Dewar, C.J., Kennedy, D. (1999). Defect and disease detection in potato tubers. *Precision Agriculture and Biological Quality*, 3543: 199-207.
- Muir, A.Y., Ross, D.W., Dewar, C.J., Kennedy, D. (1999). Defect and disease detection in potato tubers. *Proc. SPIE 3543, Precision Agriculture and Biological Quality*, 199 (January 14, 1999); doi:10.1117/12.336883.
- National Agricultural Statistics Service (NASS). (2012). Potatoes 2011 summary. United States Department of Agriculture (USDA). September 2012. ISSN: 1949-1514.
- National Potato Council, (2013). *Potato Statistical Yearbook*. National Potato Council, May 2013, Washington, D.C. USA.
- Nicolai, B.M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K.I., Lammertyn, J. (2007). Nondestructive measurement of fruit and vegetable quality by means of nir spectroscopy: a review. *Postharvest Biology and Technology*, 46: 99–118.
- Nieuwenhuizen, A.T., Tang, L., Hofstee, J.W., Müller, J., Henten, E.J.V. (2007). Colour based detection of volunteer potatoes as weeds in sugar beet fields using machine vision. *Precision Agriculture*, 8(6): 267–278.
- Noordam, J.C., Timmermans, A.J.M., Otten, G.W., Van Zwol, B.H.V. (2000). High-speed potato grading and quality inspection based on a color vision system. In J.C. Stover, K.W. Tobin (Eds.), *machine vision applications in industrial inspection* viii (pp. 206–217).
- Noordam, J.C., Van Der Broek, W.H.A.M. Buydens, L.M.C. (2004). Perspective of inline control of latent defects and diseases on French fries with multispectral imaging. In *Monitoring food safety, agriculture, and plant health.- Bellingham: International Society for Optical Engineering, (Proceedings of SPIE 5271)*.
- Nylund, R., Hemphill, E.P., Lutz, J.M. (1955). Mechanical damage to potatoes during harvesting and handling in the Red River operation in the Red River valley of Minnesota and North Dakota. *American Potato Journal*, 32(7): 237– 247.
- Nylund, R.E., Lutz, J.M. (1950). Separation of hollow heart potato tubers by means of size grading, specific gravity, and x-ray examination. *American potato journal*, 27(6): 214–222.
- Olinger, J.M., Griffiths, P.R. Burger, T. (2001). Theory of diffuse reflection in the nir region. In D.A. Burns, and E.W. Ciurczak (Eds.), *handbook of near-infrared analysis* (pp. 19-52). Marcel Dekker, Inc. New York, USA.
- Ophuis, B.G., Heszen, J.C., Kroesbergen, E. (1958). The influence of temperature during handling on the occurrence of blue discoloration inside potato tubers. *European Potato Journal*, 1(3): 48–65.

- Panigrahi, S., Wiesenborn, D., Orr, P., Schape, L., Bierwagen, G. (1996). Spectral reflectance properties of French fries. *Applied Engineering in Agriculture*, 12(6): 721-724.
- Parke, D. (1963). The resistance of potatoes to mechanical damage caused by impact loading. *Journal of Agricultural Engineering Research*, 8(3): 173–177.
- Pedreschi, F. (2009). Fried and dehydrated potato products. In J. Singh, I. Kaur (Eds.), *advances in potato chemistry and biotechnology* (pp.319-337). Academic Press, Elsevier, USA.
- Pedreschi, F., Leon, J., Mery, D., Moyano, P. (2006). Development of a computer vision system to measure the color of potato chips. *Food Research International*, 39(10): 1092–1098.
- Pedreschi, F., Mery, D., Bunger, A., Verónica, Y. (2010a). Computer vision classification of potato chips by color. *Journal of Food Process Engineering*, 34(5): 1714-1728.
- Pedreschi, F., Mery, D., Mendoza, F., Aguilera, J.M. (2004). Classification of potato chips using pattern recognition. *Journal of Food Science*, 69(6): 264-270.
- Pedreschi, F., Segtnan, V.H., Knutsen, S.H. (2010b). On-line monitoring of fat, dry matter and acrylamide contents in potato chips using near infrared interactance and visual reflectance imaging. *Food Chemistry*, 121(2): 616–620.
- Polder, G., Van der Heijden, G.W.A.M., Young, I.T. (2002). Spectral image analysis for measuring ripeness of tomatoes. ASAE Paper No: 003089. *Transactions of the ASAE*, 45(4): 1155-1161.
- Porteos, R.L., Muir, A.Y., Wastie, R.L. (1981). The identification of diseases and defects in potato tubers from measurements of optical spectral reflectance. *Journal of Agricultural Engineering Research*, 26: 151-161.
- Pringle, B., Bishop, C., Clayton, R. (2009). *Potatoes postharvest*. CAB International, Oxfordshire, UK.
- Pritchard, M.K., Scanlon, M.G. (1997). Mapping dry matter and sugars in potato tubers for prediction of whole tuber process quality. *Canadian Journal of Plant Science*, 77: 461-467
- Qiao, J., Wang, N., Ngadi, M.O., Baljinder, S. (2005). Water content and weight estimation for potatoes using hyperspectral imaging. *Proceedings the 2005 ASAE Annual International Meeting*. Tampa, Florida, USA.
- Qin, J., Lu, R. (2007). Measurement of the absorption and scattering properties of turbid liquid foods using hyperspectral imaging. *Journal of Society for Applied Spectroscopy*, 61(4), 388-396.
- Qin, J., Lu, R. (2005). Hyperspectral diffuse reflectance for determination of the optical properties of milk and fruit and vegetables juices. *Proc. SPIE 5996*. Optical Sensors and

Sensing Systems for Natural Resources and Food Safety and Quality, 59960Q (November 08, 2005); doi:10.1117/12.630691.

- Rady, A.M. Effect of mechanical treatments on handling losses of potato tubers. (2006). Master thesis. Department of Agricultural Engineering, faculty of Agriculture, Alexandria University, Alexandria, Egypt.
- Rady, A.M., Guyer, D.E. (2014). Utilization of visible/near-infrared spectroscopic and wavelength selection methods in sugar prediction and potatoes classification. *Journal of Food Measurement and Characterization*. Published online in September, 2014.
- Rady, A.M., Guyer, D.E., Kirk. W., Donis-González, I.R. (2014). The potential use of visible/near infrared spectroscopy and hyperspectral imaging to predict processing-related constituents of potatoes. *Journal of Food Engineering*, 135: 11–25.
- Rady, A.M., Soliman, N. (2013). Evaluation of surface effect, on mechanical damage of potato tubers using different methods. Annual International ASABE Meeting, Kansas City, Missouri. USA. Paper No.131621826
- Rama, M.V., Narasimham, P. (2003). Potatoes and related crops. In B. Caballero (Ed.) *encyclopedia of food science and nutrition* (pp. 4658-4680). Academic Press. USA.
- Razmjoooy, N., Mousavi, B.S., Soleymani, F. (2012). A Real-time mathematical computer method for potato inspection using machine vision. *Computers and Mathematics with Applications*, 63(1): 268–279.
- Rios-Cabrera, R., Lopez-Juarez, I., Sheng-Jen, H. (2009). ANN analysis in a vision approach for potato inspection. *Journal of Applied Research and Technology*, 6(2): 106–119.
- Romani, S., Rocculi, P., Mendoza, F., Rosa, M.D. (2009). Image characterization of potato chip appearance during frying. *Journal of Food Engineering*, 93 (4): 487–494.
- Saldaña, E., Siche, R., Luján, M., Quevedo, R. (2013). Review: computer vision applied to the inspection and quality control of fruits and vegetables. *Brazilian Journal of Food Technology*. 16(4): 254-272.
- Samanta, D., Paramita P.C., Ghosh, A. (2012). Scab diseases detection of potato using image processing. *International Journal of Computer Trends and Technology*, 3(1): 109–113.
- Sandra Segnini, S., Dejmek, P., Öste, R. (1999). A low cost video technique for colour measurement of potato chips. *Lebensm-Wiss. U.-Technol.*, (32): 216-222.
- Sarkar, N., Wolfe, R.R. (1983). Potential of ultrasonic measurements in food quality evaluation. *Transactions of the ASAE*, 26: 624–629.

- Scanlon, M.G., Pritchard, M.K., Adam, L.R. (1999). Quality evaluations of processing potatoes by near infrared reflectance. *Journal of the Science of Food and Agriculture*, 79 (5): 763-771.
- Scanlon, M.G., Roller, R., Mazza, G., Pritchard, M.K. (1994). Computerized video image analysis to quantify color of potato chips. *American Journal of Potato Research*, 71(11): 717-733.
- Schippers, P.A. (1976). The relationship between specific gravity and percentage dry matter in potato tubers. *American Potato Journal*, 53: 111-122.
- Segnini, S., Dejmek, P., Öste, R. (1999). A low cost video technique for colour measurement of potato chips. *Lebensm-Wiss. Food Science and Technology*, (32): 216-222.
- Segtnan, V.H., Kita, A., Mielnik, M., Jørgensen, Knutsen, S.H. (2006). Screening of acrylamide contents in potato crisps using process variable settings and near-infrared spectroscopy. *Molecular Nutrition and Food Research*, 50(9): 811–817.
- Shenk, J. S., Barnes, R. F. (1977a). Current Status of Infrared Reflectance. In *Proceeding of The 34th Southern Pasture and Forage Crop Improvement Conference*, Auburn, Alabama, 57-62.
- Shenk, J.S., Norris, K.H. Barnes, R.F. Fissel, G.W. (1977b). Forage and Feedstuff Analysis with Infrared Reflectance Spectro/computer System XIII International Grassland Congress, Leipzig, 1440-1441.
- Shenk, J.S., Workman, J.J., Westerhaus, M.O. (2001). Application of nir spectroscopy to agriculture products. In D.A. Burns, E.W. Ciurczak (Eds.), *handbook of near-infrared analysis* (pp.419-474). Marcel Dekker, Inc., USA.
- Shiroma, C., Rodriguez-Saona, L. (2007). Rapid quality control of potato chips using near and mid-infrared spectroscopy. Presentation, Department of Food Science and Technology. The Ohio State University, Columbus, OH, USA.
- Simmonds, N.W. (1977). Relations between specific gravity, dry matter content and starch content of potatoes. *Journal of Potato Research*, 20(2): 137-140.
- Singh, B., Wang, N., Prasher, N., Ngadi. M. (2004). A Spectroscopic technique for water content determination in potato. The 2004 ASAE/CSAE annual international meeting. Ottawa, Ontario, Canada.
- Skoog, D.A., Holler, F.J., Crouch, S.R. (2007). Components of optical instruments. In D.A. Skoog, F.J. Holler, S.R. Crouch (Eds.), *principals of instrumental analysis* (6th Ed, pp 164-214). Thomson Brooks/Cole. Canada.
- Slight, D.L. (1966). Some x-ray absorption and scatter properties of potatoes and stones. *Journal of Agricultural Engineering Research*, 11(3): 148–151.

- Sonego L., Benarie, R., Raynal, J., Pech, J.C. (1995). Biochemical and physical evaluation of textural characteristics of nectarines exhibiting woolly breakdown: nmr imaging, x-ray computed tomography and pectin composition. *Postharvest Biology and Technology*, 5(3):187–98.
- Sowokinos, J.R. (2007). Internal physiological disorders and nutritional and compositional factors that affect market quality. In D. Vreugdenhil, J. Bradshaw, C. Gebhardt, F. Govers, D.K.L. Mackerron, M.A. Taylor, H.A. Ross (Eds.), *potato biology and biotechnology advanced perspectives* (pp.501-523). Elsevier Press, UK.
- Stadler, R.H., Blank, R., Varga, N., Robert, F., Hau, J., Guy, P.A., Robert, M-C., Reidiker, S. (2002). Acrylamide from Maillard reaction products. *Nature*, 419: 449-450.
- Stark, J.C., Love, S.L. (2003). Tuber quality. In J.C. Stark, S.L. Love (Eds.), *potato production systems* (pp. 329-342), University of, Idaho, Moscow, USA.
- Stephenson, K.Q., Rotz, C.A., Singh, M. (1979). Selective sorting by resonance techniques. *Transactions of the ASABE*, 22(2): 279-282.
- Storey, R.M.J. (2007). The harvested crop. In D. Vreugdenhil (Ed.), *potato biology and biotechnology advances and perspectives* (pp. 441-470). Elsevier, UK.
- Storey, R.M.J., Davis, H.V. (1992). Tuber quality. In P. Harris (Ed.), *the potato crop* (pp. 507-569). Chapman and Hall, London, UK.
- Story, A.G. (1973). Spectral reflectance of light and infrared radiation by potatoes, stones, and soil clods. *Transactions of the ASAE*, 16(2): 302–303.
- Story, A.G., Raghavan, G.S.V. (1971). Sorting potatoes from stones and soil clods by infrared reflectance. *Transactions of the ASAE*, (16): 304–309.
- Subedi, P.P., Walsh, K.B. (2009). Assessment of potato dry matter concentration using short-wave near-infrared spectroscopy. *Potato Research*, 52 (1): 67-77.
- Tao, Y., Morrow, C.T., Heinemann, P.H., Sommer, J.H.. (1990). Automated machine vision inspection of potatoes. ASAE paper no. 90-3531.
- Tao, Y., Heinemann, P.H., Varghese, Z., Morrow, C.T. Sommer III, H.J. (1995a). Machine vision for color inspection of potatoes and apples. *Transactions of the ASAE*, 38(5): 1555–1561.
- Tao, Y., Morrow, C.T., Heinemann, P.H., Sommer III, H.J. (1995b). Fourier-based separation technique for shape grading of potatoes using machine vision. *Transactions of the ASAE*, 38(3): 949–957.
- Tarn, T.R., Tai, G.C.C., De Jong, H., Murphy, A.M., Seabrook, J.E.A. (1992). Breeding potatoes for long-day, temperature climates. *Plant Breeding Reviews*, 9: 217-232.

- Thornton, M., Bohl, W. (2000). Preventing potato bruise damage. College of Agriculture University of Idaho. Agricultural Experiment Station bul. 725(Revised).
- Thygesen, L.G., Thybo, A.K., Engelsen, S.B. (2001). Prediction of sensory texture quality of boiled potatoes from low-field 1 h nmr of raw potatoes. The role of chemical constituents. *Lebensmittel-Wissenschaft und-Technologie*, 34: 469-477.
- Thypo, A.K., Andersen, H.J., Karlsson, A.H., Dønstrup, S., Stødtkilde- Jørgensen, H. (2003). Low-field relaxation and nmr-imaging as tools in differentiation between potato sample and determination of dry matter content in potatoes. *Lebensm-Wiss. U- Technol.*, 36: 315-322.
- Thypo, A.K., Bechmann, I.E., Martens, M., Engelsen, S.B. (2000). Prediction of sensory texture of cooked potatoes using uniaxial compression, near infrared spectroscopy and low field 1h nmr spectroscopy. *Lebensm-Wiss. U- Technol.*, 33: 103-111.
- Vanoli, M., Rizzolo, A., Spinell, L., Parisi, B., Torrcelli, A. (2012). Non destructive detection of internal brown spot in potato tubers by time-resolved reflectance spectroscopy: preliminary results on a susceptible cultivar. Article No. P1370. Proceeding the International Conference of Agricultural Engineering, CIGR-Ageng2012, Valencia, Spain.
- Varmuza, K., Filzmoser, P. (2009). Introduction to multivariate statistical analysis in chemometrics (1st Edn., pp. 283-292). CRC Press, USA.
- Vibhute, A., Bodhe, S.K. (2012). Applications of image processing in agriculture: a survey. *International Journal of Computer Applications* (0975–8887), 52(2): August 2012.
- Volbracht, O., Kuhnke, U. (1956). Mechanische beschädigungen an kartoffe In (mechanical injuries of the potatoes). *Der Kartoffel bau heft* No 4 and 5. April and May. Translation No.125, Niae, Silose , Bedford, England.
- Walsh, K.B., Golic, M., Greensill, C.V. (2004). Sorting of fruit using near infrared spectroscopy: application to a range of fruit and vegetables for soluble solids and dry matter content. *Journal of Near Infrared Spectroscopy*, 12(3): 141-148.
- Wang, N., El Masry, G. (2010). Bruise detection of apples using hyperspectral imaging. In D.-W. Sun (Ed.), *hyperspectral imaging for food quality analysis and control* (pp.295-320). Elsevier Press, UK.
- Watts, K.C., Russel, L.T. (1985). A review of techniques for detecting hollow heart in potatoes. *Canadian Society for Bioengineering*, 27: 85-90.
- Williams, P. (2007). Near-infrared technology – getting the best out of light (5th Ed.). PDK Grain, Nanaimo, Canada. A short course in the practical implementation of near-infrared spectroscopy for the user.

- Willson, J.H., Lindsay, A.M. (1969). The relation between specific gravity and dry matter content of potato tubers. *American Potato Journal*, 46: 323-328.
- Wise, B.M., Gallagher, N.B., Bro, R., Shaver, J.M., Windig, W., Kock, R.S. (2006). *PLS_Toolbox 4.0 for use with Matlab* (pp. 137-192). Eigenvector Research, Inc. WA, USA.
- Woodbury, G.W., Weinheimer, W.H. (1964). Specific gravity- solids correlations in Russet Burbank with respect to point of origin and storage history. *American Potato Journal*, 42: 98-104.
- Woolfe, J.A. (1987). *The Potato in human diet*, Cambridge University press. UK.
- Workman, J., Weyer, L. (2008). Water. In J. Workman, L. Weyer (Eds.), *practical guide to interpretive near-infrared spectroscopy* (pp. 63-70). CRC Press, FL, USA.
- Yaptenco, K.F., Kawakamis, S., Takano, K. (2000). Nondestructive determination of sugar content in ‘Danshaku’ potato (*solanum tuberosum* l.) by near infrared spectroscopy. *Journal of Agricultural Science, Tokyo Nogyo Daigaku*, 44(4): 284-294.
- Yin, H., Panigrahi, S. (2004). Image processing techniques for internal texture evaluation of French fries. *Applied engineering in agriculture*, 20(6): 803–811.
- Zahara, M., Mclean, J.G., Wright, D.N. (1961). Mechanical injury to potato tubers during harvesting. *California Agriculture*. August, 1961.
- Zhou, L., Vikram C., Yongmin K. (1998). PC-based machine vision system for real-time computer-aided potato inspection. *International Journal of Imaging Systems and Technology*, 9(6): 423–433.
- Zyzak, D.V., Sanders, R.A., Stojanovic, M., Tallmadge, D.H., Eberhart, B.L., Ewald, D.K., Gruber, D.C., Morsch, T.R., Strothers, M.A., Rizzi, G.P., Villagran, M.D. (2003). acrylamide formation mechanism in heated foods. *Journal of Agriculture and Food Chemistry*, 51: 4782-4787.