NEAR INFRARED SPECTROSCOPIC INVESTIGATION OF LIPID OXIDATION IN MODEL SOLID FOOD SYSTEMS

By

CHRISTOPHER G. IZZO

A Dissertation submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Food Science

Written under the direction of

Professor Karen M. Schaich

And approved by

_____

_____

_____

_____

New Brunswick, New Jersey

May 2019

ABSTRACT OF THE DISSERTATION

NEAR INFRARED SPECTROSCOPIC INVESTIGATION OF LIPID OXIDATION IN MODEL SOLID FOOD SYSTEMS

By

CHRISTOPHER G. IZZO

Dissertation Director:

Professor Karen M. Schaich

Reports of the application of near infrared (NIR) spectroscopy to analyses of lipid oxidation in solid foods generally indicate poor performance. To elucidate reasons for this, effects of sample packing and presentation (off-centered rotation) on NIR analyses were examined in a sampling system miniaturized to employ amounts of material feasible for research studies. Packing and presentation conditions affording the best performance in qualitative studies were utilized in quantitative assays to determine the ability of NIR to monitor lipid oxidation in model solid food systems by comparison with reference chemical analyses of conjugated dienes, lipid hydroperoxides, and carbonyl products.

Preliminary investigation indicated constant forming pressure and rotational averaging during scanning reduced variation among replicate scans of mixtures of up to 15% (w:w) lipid with white rice flour. Neat pecan or canola oils oxidized at 40°C for up to sixteen weeks and assayed chemically for conjugated dienes, lipid hydroperoxides and carbonyls were used to prepare 7.5% (w/w) oil : white rice flour samples for NIR analysis with constant pressure and rotation. Canola oxidized more readily than pecan oil, reaching apparent maxima for conjugated dienes and peroxides; however, carbonyls developed only near the end of incubation.

NIR models of oxidation used either the full spectrum (4000 - 10,000 cm^{-1}) or wavenumber ranges selected by statistical model improvement techniques. Full spectrum models of conjugated dienes or peroxides for pecan oil samples showed very poor correlations with chemical analyses; neither was improved by wavenumber selection. Full spectrum models for canola oil samples were slightly better and improved with wavenumber selection.

Peroxide value model quality rose with sample numbers; the opposite occurred for conjugated dienes. The best peroxide value models included far fewer wavenumbers than conjugated diene models, which were more susceptible to interference from various sources. Results from wavenumber selection appeared pathway dependent, varying with samples used and pretreatments applied in the initial model. Spectral reproducibility among nominally identical samples was the primary hindrance to quantitative correlations for conjugated dienes and peroxide values. Thus, improvements in sample presentation mechanisms and software may render NIR suitable for quantitative analysis of lipid oxidation in solid food systems.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my major advisor, Dr. Karen Schaich, for her guidance, assistance, encouragement and friendship over the course of this project. I would also like to thank her as well as Dr. Donald Schaffner for their steadfast support.

I thank the members of my committee, Dr. Thomas Hartman, Dr. Chaim Frenkel and Dr. Henrik Pedersen for their guidance, helpful discussions and thoughtful review of this work and the proposal which led to it.

I am grateful to the faculty and staff of the Department of Food Science for their help and friendship. I thank Dr. Thomas Hartman, Dr. Mukund Karwe, and Dr. Kit Yam for sharing their equipment. I also thank Joseph Lech for sharing his analytical expertise. I owe special thanks to David Petrenka, William Sumal and Frank Caira for keeping the equipment running and for creating all the critical machines and devices that didn't come from the factory. I also thank Laura Amador, Deborah Koch, Karen Ratzan, Yakov Uchitel, Rosanne Vaccaro and Irene Weston for their administrative support.

I am profoundly grateful to Dr. Barbara Zilinskas for her counsel and friendship. Her loan of equipment from the beginning of this project was as invaluable as it was generous.

I thank Ryanne Palermo and Mark Sullivan at Buchi for their guidance and assistance. I also thank the Kinloch Plantation Company for their generous donation of fresh pecan oil.

I thank my friends and fellow students Xin Tian, Jia Xie, Brandon Bogusz, Emily Nering, Karen Chang, Lynn Yao, Teng Peng, Weiyue Wang, Ed Pappas, Morgan Kandrac and Zachary Yeager for their assistance and collaboration.

I thank all my friends and colleagues at Rutgers and UMDNJ for their encouragement, counsel and support.

Most importantly, I thank my family for their love, patience, guidance, help and encouragement throughout this and all of my endeavors. I thank God for each of them and for all the blessings in my life.

TABLE OF CONTENTS

ABSTRACT OF THE DISSERTATION ii

ACKNOWLEDGEMENTS iv

TABLE OF CONTENTS v

LIST OF ABBREVIATIONS x

LIST OF TABLES xi

LIST OF ILLUSTRATIONS xiv

1. INTRODUCTION 1

2. BACKGROUND 3

2.1 NIR Basics 3

2.2 General Procedures for the Use of Chemometrics in Spectroscopy 5

2.2.1 Preparation of Spectra for Chemometric Analysis via Pretreatment 5

2.2.1.1 Normalization 5

2.2.1.2 Smoothing 7

2.2.1.3 Derivatization 9

2.2.2 Vector Representation of Spectra and Linear Transformation Using Matrices 10

2.2.3 Results of Linear Transformation: Latent Variable Isolation 13

2.2.3.1 Loadings 13

2.2.3.2 Scores 13

2.3 Limitations in the Application of NIR to Food Analysis 15

2.4 Chemometric Approaches to Construct and Assess NIR Models 17

2.5 Review of the Literature: NIR Studies of Lipid Oxidation 17

2.5.1 Lipid Hydroperoxides 17

2.5.1.1 NIR Analyses in Neat Oils 17

2.5.1.1.1 Poor to Fair Quality NIR Models 19

2.5.1.1.2 Improved Quality NIR Models from ASL Studies 22

2.5.1.2 NIR Analyses in Complex Food Systems 28
2.5.2 Conjugated Dienes 32
2.5.3 Carbonyls 37
2.5.3.1 Total Carbonyl Value 37
2.5.3.2 Hexanal 37
2.5.3.3 Free Fatty Acids 38
2.6 Issues raised by analysis of the literature 40
3. HYPOTHESIS AND OBJECTIVES 41
3.1 Hypothesis and Overall Objective 41
3.2 Specific Objectives 41
4. EXPERIMENTAL PROCEDURES 43
4.1 Experimental Design 43
4.1.1 Development of a Miniaturized FT-NIR Sample Platform 43
4.1.2 Assessment of FT-NIR for Quantitation of Indicators of Lipid Oxidation 44
4.2 Materials 47
4.2.1 Solvents 47
4.2.2 Reagents 47
4.2.3 Sample Materials 48
4.3 Equipment 48
4.4 Experimental Methods 55
4.4.1 Discrimination of 15% w/w Lipid : White Rice Flour Mixtures 55
4.4.2 Discrimination of Degree of Lipid Oxidation : Qualitative ASL Study 56
4.4.3 Discrimination of Identical Samples Subjected to Different Pressures 56
4.4.4 Effect of Sample Rotation 57
4.4.4.1.1 Replicate Scans of a Single Sample 58
4.4.4.1.2 Replicate Scans of Multiple Samples 59

4.4.5 Accelerated Shelf Life Study: Quantitative Analysis of Lipid Oxidation 59
4.4.5.1 Conjugated Dienes 60
4.4.5.2 Lipid Hydroperoxides 61
4.4.5.3 Carbonyls 62
4.4.5.3.1 Recrystallization of DNPH 63
4.4.5.3.2 Preparation of DNPH Reagent 63
4.4.5.3.3 DNPH Reaction and HPLC Assay 64
4.4.5.4 Near Infrared Spectroscopy 65
4.4.5.4.1 Data Collection 65
4.4.5.4.2 Assessment of Wavenumbers for Use in Models 66
4.4.5.4.2.1 Comparison of Spectral Variation Among and Within Groups: e 66
4.4.5.4.2.2 Cross Validation Regression Coefficients t-Test 67
4.4.5.4.2.3 Model Improvement by Wavenumber Selection 71
5. RESULTS & DISCUSSION 72
5.1 Preliminary Experiments using Near Infrared Spectroscopy 72
5.1.1 Discrimination of Fat or Oil Type in Lipid : White Rice Flour Mixtures 72
5.1.2 Discrimination of Oxidation Status in Lipid : White Rice Flour Mixtures 75
5.1.3 Effect of Sample Packing on Variability of NIR Response 85
5.1.3.1 15% Canola Oil: White Rice Flour Samples 85
5.1.3.2 7.5% Canola Oil: White Rice Flour Samples 87
5.1.4 Effect of Sample Rotation on Variability in NIR Spectra 91
5.1.4.1 Variability Among Replicate Scans of a Single Sample 91
5.1.4.2 Variability Among Replicate Scans of Multiple Samples 95
5.2 Quantitative Analysis of Lipid Oxidation: Accelerated Shelf Life Study 99
5.2.1 Experimental Results 99
5.2.1.1 Conjugated Dienes Reference Assay 99

5.2.1.2 Lipid Hydroperoxides Reference Assay 101
5.2.1.3 Non-Volatile Carbonyl Products Reference Assay 102
5.2.1.4 Near Infrared Spectroscopy 109
5.2.2 Full Spectrum Predictive NIR Models for Lipid Oxidation Parameters 109
5.2.2.1 NIR Assays of Conjugated Dienes (Oxidation Index) 112
5.2.2.1.1 Pecan Oil Samples 112
5.2.2.1.2 Canola Oil Samples 114
5.2.2.1.3 Combined Samples 115
5.2.2.2 NIR Assays of Lipid Hydroperoxides 117
5.2.2.2.1 Pecan Oil Samples 117
5.2.2.2.2 Canola Oil Samples 119
5.2.2.2.3 Combined Samples 120
5.2.3 Predictive NIR Models for Lipid Oxidation using Selected Wavenumbers 123
5.2.4 NIR Model Improvement: Wavenumber Selection using the e Coefficient 126
5.2.4.1 Conjugated Diene Model Improvement Based on e Values 127
5.2.4.2 Conjugated Diene Model Improvement Based on the Numerator of e 130
5.2.4.3 Peroxide Value Model Improvement Based on e Values 132
5.2.4.4 Peroxide Value Model Improvement Based on the Numerator of e 134
5.2.5 NIR Model Improvement: Cross Validation Regression Coefficients t Test 137
5.2.5.1 Conjugated Diene Model Improvement by Significance Testing of t 137
5.2.5.2 Conjugated Diene Model Improvement by Ranking of t 141
5.2.5.3 Peroxide Value Model Improvement by Significance Testing of t 144
5.2.5.4 Peroxide Value Model Improvement by Ranking of t 146
5.2.6 Summary of Best Models 149
5.2.6.1 Conjugated Dienes 152
5.2.6.2 Peroxide Values 155

5.2.6.3 Interpretation of Wavenumbers 160

5.2.6.3.1 Models of Conjugated Dienes 160

5.2.6.3.2 Models of Peroxide Values 168

5.2.6.4 Summary of Tentative Functional Group Assignments in Models 176

6. SUMMARY AND CONCLUSION 179

7. FUTURE WORK 186

8. REFERENCES 188

APPENDIX A: Colthup Chart from NIR Studies of Lipids and Lipid Oxidation [Adapted from NIRCal 5.4 Software Module, Appendix 4a of Workman & Weyer (2008) and Lipid Oxidation Literature] 193

APPENDIX B: Statistical Measures of Model Fitness 213

B.1 Defining Error to Assess Model Fitness 213

B.2 Coefficients of Correlation (r) and Determination (r^2) 215

B.3 Q-Value 216

LIST OF ABBREVIATIONS

ACN	Acetonitrile
ASL	Accelerated Shelf Life (Study)
CD	Conjugated Dienes
FT-NIR	Fourier Transform Near Infrared (Spectroscopy)
IPA	Isopropyl Alcohol
MLR	Multiple Linear Regression
NIR	Near Infrared
OI	(Klein's) Oxidation Index
PC	Principal Component
PCA	Principal Components Analysis
PCR	Principal Components Regression
PLS	Partial Least Squares (Analysis)
PLSR	Partial Least Squares Regression
PV	Peroxide Value
RMSEC	Root Mean Square Error of Calibration
RMSECV	Root Mean Square Error of Cross Validation
RMSEP	Root Mean Square Error of Prediction
RPD	Residual Predictive Deviation
$\mathbf{SD_{Reference}}$	Standard Deviation of Values from the Reference Assay
SDD	Standard Deviation of Differences
SEC	Standard Error of Calibration
SECV	Standard Error of Cross Validation
SEP	Standard Error of Prediction
SNV	Standard Normal Variate (Normalization)
TE	Trolox Equivalents

LIST OF TABLES

Table 1. Windows for smoothing functions. 8

Table 2. Poor to fair quality NIR models for peroxide values in lipids from the literature. 21

Table 3. Partial Least Squares (PLS) NIR models for peroxide values in lipids from the literature. 25

Table 4. Single and double wavelength NIR models from MLR for peroxide values in lipids from the literature. 26

Table 5. Partial Least Squares (PLS) NIR models for peroxide values in solid food systems from the literature. 33

Table 6. NIR models for conjugated dienes in lipids from the literature. 36

Table 7. NIR models of hexanal content from the literature. 39

Table 8. HPLC Gradient Conditions for DNPH Carbonyl Assay. 64

Table 9. Q-value criteria of qualitative model for discrimination of 15% lipid : white rice flour mixtures. 73

Table 10. Q-value criteria of qualitative model for discrimination of oxidation status of soybean oil in 15% soybean oil : white rice flour mixtures. 77

Table 11. Q-value criteria of qualitative model for discrimination of oxidation status of canola oil in 15% canola oil : white rice flour mixtures 78

Table 12. Possible chemical assignments from 15% canola oil oxidation status model loadings: (a) Principal component 1; (b) Principal component 2. 81

Table 13. Q-value criteria of qualitative model for pressure discrimination in 7.5% canola oil : white rice flour samples. 88

Table 14. Eigenvalues of qualitative model for pressure discrimination in 7.5% canola oil : white rice flour samples. 89

Table 15. Standard deviation of scores of one 10% canola oil : white rice flour sample. 93

Table 16. Standard deviation of scores of ten 10% canola oil : white rice flour samples. 97

Table 17. Typical fatty acid composition of pecan and canola oils. 107

Table 18. Q-values for models of conjugated dienes in 7.5% pecan oil samples. 113

Table 19. Q-values for models of conjugated dienes in 7.5% canola oil samples. 114

Table 20. Best individual oil conjugated diene models for each sample set. 115

Table 21. Combined data models of conjugated dienes using SNV and first derivative pretreatment ranked by Q-value..... 116

Table 22. Combined data models of conjugated dienes using SNV and smoothing pretreatment ranked by Q-value. 117

Table 23. Q-values for models of peroxide values in 7.5% pecan oil samples. 118

Table 24. Q-values for models of peroxide values in 7.5% canola oil samples. 120

Table 25. Best individual oil peroxide value models for each sample set..... 121

Table 26. Combined data models of peroxide values using SNV and first derivative pretreatment ranked by Q-value. 122

Table 27. Combined data models of peroxide values using SNV and smoothing pretreatment ranked by Q-value. 123

Table 28. Conjugated diene models from significance testing of cross validation regression coefficient t test values at 99.5% confidence. 137

Table 29. Conjugated diene models from ranking of cross validation regression coefficient t test values. 141

Table 30. Peroxide value models from significance testing of cross validation regression coefficient t test values at 99.5% confidence. 144

Table 31. Peroxide value models from ranking of cross validation regression coefficient t test values. 146

Table 32. Statistics of best models for conjugated dienes. 153

Table 33. Statistics of best models for peroxide values. 156

Table 34. CH bands reported near key wavelengths in best conjugated diene model. 163

Table B.1. Interpretation of residual predictive deviation (RPD) [From Williams, 2001]. 216

Table B.2. Q-value criteria for qualitative models [From NIRCal 5.4 Software Manual]..... 217

Table B.3. Q-value criteria for quantitative models [From NIRCal 5.4 Software Manual]..... 217

LIST OF ILLUSTRATIONS

Figure 1. Effect of normalization on spectra. 6

Figure 2. Effect of smoothing on spectra. 9

Figure 3. Effect of derivatization on spectra. 10

Figure 4. Effect of linear transformation on the vector representation of a spectrum. 11

Figure 5. Linear transformation applied to a set of samples. 12

Figure 6. Loadings plot. 14

Figure 7. Scores plot. 14

Figure 8. Resolution of samples with different relative variation among replicates. 16

Figure 9. Overview of techniques for chemometric modeling of spectral data. 18

Figure 10. Experimental design for preliminary investigation of FT-NIR analysis of model solid food systems. 45

Figure 11. Experimental design for accelerated shelf life studies of pecan and canola oils. 46

Figure 12. Buchi NIRFlex Solids N500 FT-NIR instrument used in this study. 49

Figure 13. Carver hydraulic press used to prepare sample tablets for NIR analysis. 51

Figure 14. Lever press constructed for compacting samples at discrete pressures. 51

Figure 15. Compacted 7.5% (w/w) canola oil : white rice flour samples. 52

Figure 16. Schematic drawing (left) and photo (right) of rotating sample holder. 53

Figure 17. Insertion of compacted sample on borosilicate glass disc in sample holder. 54

Figure 18. Scanning of a stationary sample tablet on a washer. 55

Figure 19. Cross validation of models created from a set of ten samples. [Adapted from Williams (2001) and Naes et al. (2002)]. 68

Figure 20. Cross validation regression coefficients t-test for a model created from a set of ten samples. [Adapted from Martens and Martens (2000), Tukey (1958) and NIRCal 5.4 Software Manual]. 70

Figure 21. Scores plot of qualitative model for discrimination of 15% lipid : white rice flour mixtures 73

Figure 22. Pretreated spectra of different fats or oils in 15% lipid : white rice flour mixtures from 5760-5900 cm^{-1}. 74

Figure 23. Scores plot of discrimination of oxidation status of soybean oil in 15% soybean oil : white rice flour mixtures. 76

Figure 24. Scores plot of discrimination of oxidation status of canola oil in 15% canola oil : white rice flour mixtures 77

Figure 25. Loadings plots for 15% canola oxidation status model: (a) Principal component 1; (b) Principal component 2. 80

Figure 26. 15% Canola oil oxidation status model scores: (a) T0; (b) 1 Week; (c) 2 Weeks; and (d) 3 Weeks incubation.. 83

Figure 27. Scores plot showing effects of pressure on variability in 15% canola oil : white rice flour samples using all four forming pressures assayed. 86

Figure 28. Scores plot showing effects of pressure on variability in 15% canola oil : white rice flour samples using the lowest three forming pressures assayed. 86

Figure 29. Scores plot of pressure discrimination model in 7.5% canola oil : white rice flour samples using all four forming pressures assayed.. 88

Figure 30. Loadings plots for 7.5% canola oil pressure discrimination model: (a) Principal component 1; (b) Principal component 2; (c) Principal component 3. 90

Figure 31. Scores plots of a 10% canola oil: white rice flour sample: (a) Stationary; (b) Rotating. 92

Figure 32. Standard deviation among scores of 21 scans of a single 10% canola oil : white rice flour sample. 93

Figure 33. Expanded view of scores plot from scans of a single pressed 10% canola oil : white rice flour sample acquired with rotation. 94

Figure 34. Expanded view of scores plot from scans of a single powdered 10% canola oil : white rice flour sample acquired with rotation 94

Figure 35. Scores plots of 10% canola oil: white rice flour samples: (a) Stationary; (b) Rotating. 96

Figure 36. Standard deviation among scores of triplicate scans of ten 10% canola oil : white rice flour sample tablets. 97

Figure 37. Expanded view of scores plot of rotating sample tablets pressed at 415 psi. 98

Figure 38. Box and whiskers plots showing distribution of oxidation index values for oils during the 40°C ASL study. 99

Figure 39. Oxidation index (conjugated dienes) for (a) pecan and (b) canola oils incubated at 40°C. 100

Figure 40. Box and whiskers plots showing distribution of peroxide values (meq/Kg) for oils during the 40°C ASL study. 102

Figure 41. Peroxide values for (a) pecan and (b) canola oils incubated at 40°C.. 103

Figure 42. 360 nm chromatograms following DNPH derivatization of canola oil 104

Figure 43. 233 nm chromatograms of oils incubated at 40°C following DNPH derivatization. 106

Figure 44. Original NIR reflectance spectra of 7.5% (w/w) nominal mixtures of oils from the ASL study with white rice flour: (a) Pecan oil samples and (b) Canola oil samples. 110

Figure 45. Normalized and smoothed average NIR spectra for each of the twenty one 7.5% oil : white rice flour batches from ASL studies of (a) Pecan oil and (b) Canola oil. 111

Figure 46. Comparison of regression coefficient statistics for two initial models of peroxide values made from different sample sets. 125

Figure 47. Plots of (a) e coefficient values; and (b) the numerator and denominator of e. 128

Figure 48. Refinement of conjugated dienes model by wavenumber selection based on e. 129

Figure 49. Wavenumber ranges used in models after application of e coefficient cutoffs. 129

Figure 50. Refinement of conjugated dienes model by selection of wavenumbers based on the numerator of e. 131

Figure 51. Wavenumber ranges used in models after application of numerator of e cutoffs..... 131

Figure 52. Refinement of peroxide value model by wavenumber selection based on e. 132

Figure 53. Wavenumber ranges used in models after application of e coefficient cutoffs. 133

Figure 54. Refinement of peroxide value model by selection of wavenumbers based on the numerator of e. 135

Figure 55. Wavenumber ranges used in models after application of numerator of e cutoffs..... 135

Figure 56. Q-values for models from refined numerator of e cutoffs. 136

Figure 57. Wavenumber ranges for refined numerator of e cutoffs. 136

Figure 58. Regression coefficient statistics for best initial conjugated dienes model..... 138

Figure 59. Regression coefficient statistics for initial peroxide values model subjected to improvement procedures..... 139

Figure 60. Regression coefficient statistics for conjugated dienes model after second iteration of significance testing (99.5% Confidence). 140

Figure 61. Conjugated dienes model Q-values: wavenumber selection based on cross validation regression coefficients t test cutoff thresholds for the best initial model, best first iteration model and best second iteration model. 142

Figure 62. Regression coefficient statistics for conjugated dienes model after second iteration of regression coefficient t test ranking.. 143

Figure 63. Regression coefficient statistics for peroxide value model after fourth iteration of significance testing (99.5% confidence). 145

Figure 64. Peroxide value model Q-values: wavenumber selection based on cross validation regression coefficients t test cutoff thresholds for the initial model, best first iteration model and best second iteration model..... 147

Figure 65. Regression coefficient statistics for best peroxide value model after third iteration of regression coefficient t test ranking. 148

Figure 66. Summary of wavenumber ranges isolated by techniques to improve models of conjugated dienes. 150

Figure 67. Summary of wavenumber ranges isolated by techniques to improve models of peroxide values. 151

Figure 68. Original vs. NIR predicted values of conjugated dienes (oxidation index) for best models of pecan oil samples, canola oil samples and combined sample set. 154

Figure 69. Original vs. NIR predicted peroxide values (meq / Kg) for best models of pecan oil samples, canola oil samples and combined sample set. 157

Figure 70. Characteristic wavenumbers important in lipid oxidation assays. 161

Figure 71. Wavenumber ranges used in the best models for lipid oxidation parameters. 161

Figure 72. Regression coefficients of best conjugated diene models from 7.5% oil: white rice flour samples 162

Figure 73. Assignment of functional groups to important wavenumbers of conjugated diene models.. 164

Figure 74. Regression coefficients of best peroxide value models from 7.5% oil: white rice flour samples 169

Figure 75. Assignment of functional groups to important wavenumbers of peroxide value models.. 170

Figure 76. Effect of borosilicate glass substrate on pretreated difference spectra of pecan samples scanned with and without a glass substrate in the path of the NIR beam. 175

Figure 77. Summary of tentative functional group assignments for important wavenumbers from the best models for conjugated dienes from 7.5% oil: white rice flour samples 177

Figure 78. Summary of tentative functional group assignments for important wavenumbers from the best models for peroxide values from 7.5% oil: white rice flour samples.. 178

Figure B.1. Correlation of NIR predictions with reference assay values: (a) Ideal case and (b) Real case.....214

Figure B.2. Distributions of original data and residuals from plot of NIR predicted values vs reference assay values (Figure B.1b). 215

1. INTRODUCTION

Near infrared (NIR) spectroscopy is a rapid analytical technique which if successful can supplant some traditional chemical assays and the associated time, labor and material costs. It is environmentally sound as NIR obviates the use of hazardous chemicals. Its speed can expand the scope of analytical programs at negligible cost, affording certainty of ingredient quality prior to use in manufacturing. This is particularly advantageous in an increasingly demanding regulatory environment, ensuring legal compliance and minimizing losses from producing non-salable goods.

While in theory this technology appears promising, its practical application has often been limited to simple systems or to analytes present as bulk constituents in more complex ones. NIR has been used in pharmaceutical analysis, where strict tolerances for the purity and identity of products provide uniformity advantageous to the application of the technique. It has also been used in the analysis of agricultural commodities for qualitative identification of origin as well as quantitative analysis of bulk constituents. Both pharmaceuticals and agricultural commodities are simpler analytical systems than those of food products, which are nonhomogeneous in nature. The application of NIR to food systems has met with mixed results depending on the complexity of the system and the amount of analyte present therein.

Our interest lies in the analysis of lipid oxidation in foods. In addition to the advantages discussed above, the speed of NIR could greatly benefit the determination of lipid oxidation products. These analytes pose challenges for traditional chemical assays given their transient and labile nature, particularly in solid food systems where extractions must be performed without changing the amount and distribution of lipid oxidation products. Unfortunately, the literature generally indicates poor performance among NIR analyses of lipid oxidation in solid food systems. This result reflects the fact that oxidation products are typically present in small amounts in a complex and nonhomogeneous food matrix.

To overcome the problem posed by nonhomogeneous systems, NIR practitioners have adopted averaging techniques in which each spectrum is a composite of spectra acquired from multiple locations of a sample. Given the number of replicates required to construct a reliable NIR model, such techniques can demand amounts of sample impractical for a research lab. While it is possible to reuse samples when quantities are limited, the practice involves physically disrupting the sample between spectral measurements to generate multiple views of the same sample. Such approaches are impractical for labile analytes such as lipid oxidation products.

The primary objective of this research was to assess the ability of NIR to monitor lipid oxidation in solid foods using model systems and to determine the effects of various factors on the analysis. Each model system included minor amounts (5 – 15% (w/w)) of a fat or oil mixed with white rice flour. Studies of oxidative degradation were performed by incubating the lipid constituent alone prior to mixing with flour to obtain the NIR sample. Reference chemical analyses were thus able to be performed on the lipid component without the complicating effects of a prior extraction.

Specific objectives of this research included assessing the feasibility of a miniaturized NIR sampling system to use amounts of sample manageable for a typical research lab and assessing techniques for sample handling and scanning to minimize sources of spectral variation. The ability of NIR to discriminate among model systems comprising minor amounts of (a) different lipids or (b) identical fats or oils at different stages of oxidation were investigated. We also assessed quantitative NIR assays for three specific markers of lipid oxidation.

The knowledge gained from this research will be particularly useful in rapid assays of lipid oxidation and more generally for rapid analysis of minor constituents in solid food systems. If the adverse impact of sample handling and presentation methods, sample nonhomogeneity and other factors can be identified and corrected, it may be possible to extend the cost-effective, green benefits of NIR to a much broader range of analytes in solid foods

2. BACKGROUND

2.1 NIR Basics

Near Infrared (NIR) spectroscopy has attracted considerable attention since it was first used in the analysis of agricultural commodities decades ago. Proponents of the technology assert numerous advantages including rapid sample measurement, often on the scale of seconds to minutes, minimal to no need for sample preparation, the potential for measurement without destruction of the sample, and the ability to use inexpensive optical accessories such as glass or even plastic during measurement (Osborne & Fearn, 1986). The NIR region of the spectrum spans the wavelength range from 780 to 2500 nm, or in the units commonly employed by practitioners, the wavenumbers from 4000 to 12,821 cm^{-1}. Nestled between the visible region (wavelengths lower than 780 nm) and the mid-infrared (mid-IR) (wavelengths above 2500 nm), the NIR is characterized by signals arising from overtones and combination bands of the fundamental molecular vibrations which occur in the mid-IR region (Workman & Weyer, 2008). Unlike the discrete, sharp bands afforded by these fundamental stretching, bending or deformation vibrations, their overtones and combinations are broad bands with reduced intensity which typically overlap each other (Osborne & Fearn, 1986). Thus, the NIR was for some time thought to be lacking as an analytical tool despite the recognition of utility of the mid-IR region.

NIR attracted renewed attention with the advent of modern computing, when developments in speed and processing power made the use of multivariate statistical analyses to analyze and extract information from raw spectra a reality. A wide and varied host of such analyses are encompassed within the discipline known as Chemometrics, which sits at the interface of analytical chemistry, statistics and computing (Brereton, 2003). Whereas wet chemical analysis of a sample requires either a technique capable of specifically detecting an analyte within a sample or a physical separation of the analyte from interferents prior to detection, Chemometrics offers the potential for mathematical isolation of the signal of an analyte of interest

from a complex sample using little or no sample preparation. Early NIR practitioners saw the potential of Chemometrics to obviate the perceived shortcoming of the broad, overlapping bands in this region. Moreover, the diminished signal intensity thought to render NIR of secondary import to mid-IR gave rise to a notable advantage. Given the strong signals of many materials in the mid-IR, lenses and other optical accessories used for sample presentation to the instrument are limited to expensive materials such as quartz which afford minimal spectral interferences. A host of other, more cost-effective accessories become available for use in NIR considering the reduced intensity of interfering signals from materials such as borosilicate glass and the ability to parse spectral components with Chemometrics (Osborne & Fearn, 1986).

The union of NIR analysis and Chemometrics holds significant promise for faster, greener, more cost-effective analyses when properly employed in a sampling program. These techniques are often used to supplant wet chemical techniques requiring substantially more labor and time in the preparation and performance of an analysis, as well as the use of consumables such as hazardous chemicals with their own acquisition and disposal costs. In other cases, the speed of NIR enables the performance of analyses that were not possible prior to its use, such as routine and rapid lot testing upon receipt of shipments. These analyses can be used to flag potentially problematic materials for further assessment by wet chemical techniques. This minimizes the time, labor, cost, and waste associated with such analyses as they are used only when necessary. Moreover, with little additional cost companies obtain improved certainty in the quality of materials to be used in making their products prior to manufacturing. This obviates losses incurred in using labor, materials and equipment time to produce non-salable goods.

NIR can be applied to determine constituents and/or physical properties of a sample for qualitative or quantitative purposes. Determinations of constituents have evolved with the technology. Early adopters used NIR to analyze bulk quantities such as protein, lipid, or moisture content, while more refined assays for specific compounds have since been added to the growing body of literature in the field. As a physical technique, NIR has been used in the determination of

properties such as hardness in pharmaceutical tablets (Guo et al., 1999) as well as kernel hardness in wheat and digestibility in forages (Williams, 2001). Qualitative assays categorize known samples by their spectral features to develop models for accurate identification of unknowns. An example would be to identify agricultural commodities emanating from certain growing regions. Quantitative assays correlate spectral features with known values of quantities or parameters of interest measured by reference techniques to supplant those techniques in future analyses. An example would be to monitor hexanal content in oxidized oil based on a model developed with data from GC analyses (Jensen et al., 2001).

2.2 General Procedures for the Use of Chemometrics in Spectroscopy

2.2.1 Preparation of Spectra for Chemometric Analysis via Pretreatment

It is often helpful and thus a common practice to pretreat spectra prior to the application of chemometrics. Pretreatment assists in separating the overlapping bands of the signal and minimizes the effects of noise. There are three categories of pretreatments, namely, normalization, smoothing, and derivatization. Pretreatments from multiple categories are often used, and hybrid approaches spanning categories are also possible.

2.2.1.1 Normalization

Normalization functions are used to reduce baseline variations among different spectra (NIRCal 5.4 Software Manual). The effect of normalization on a set of spectra of mixtures of oil and flour is shown in Figure 1. The 135 spectra shown include all three replicate scans acquired for each of fifteen sample tablets made from each of the three batches of oil drawn at a single time point of the quantitative accelerated shelf life study performed herein.

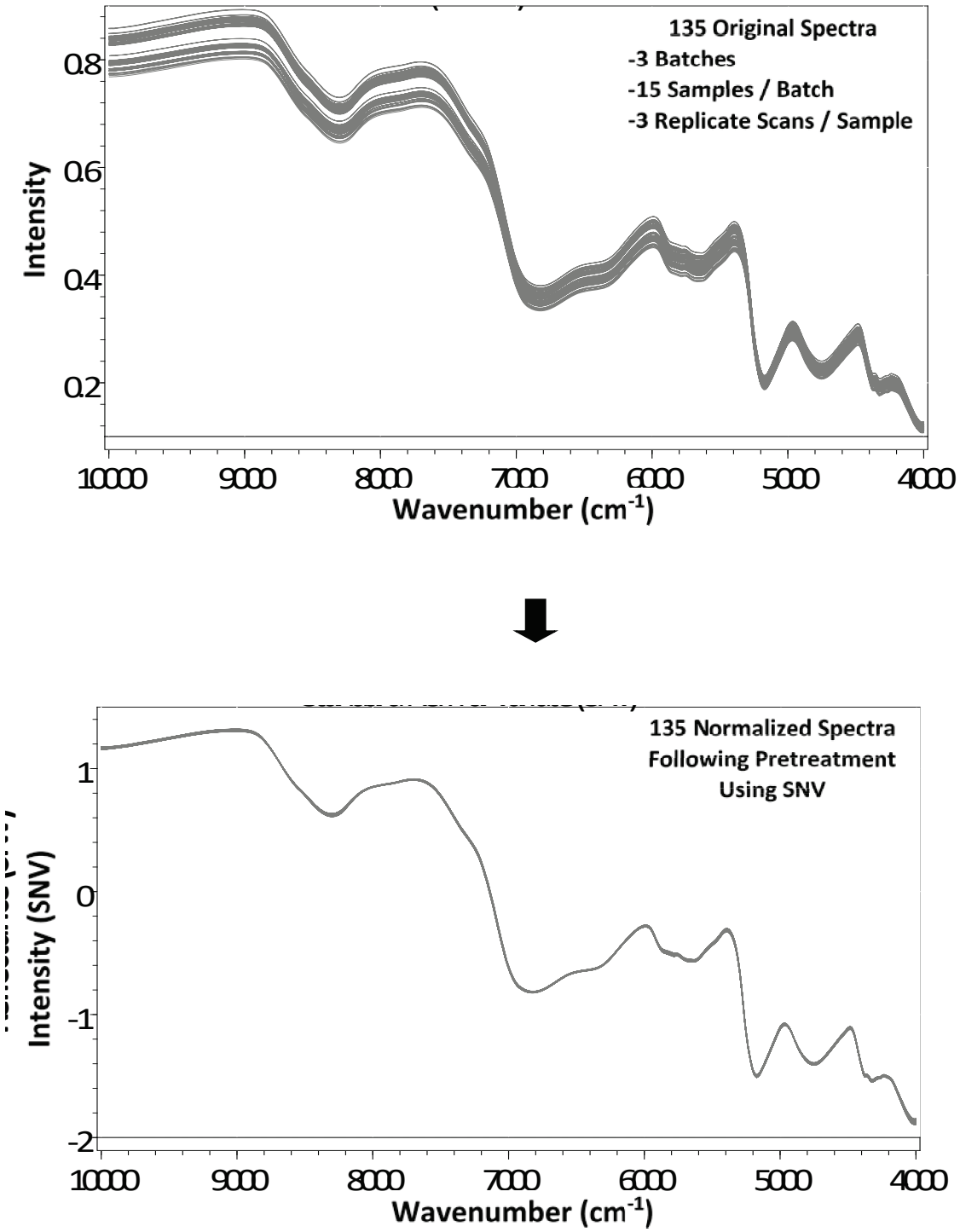


Figure 1. Effect of normalization on spectra.

In the analysis of solid systems, particularly those of granules or powders, particle size effects, packing (pressure) differences and sample thickness give rise to light scattering that confound spectra (NIRCal 5.4 Software Manual). Scattering increases with the number of interactions between light and the sample (Naes et al., 2002). This makes FT-NIR of powdered samples, which have numerous surface interactions, particularly problematic. As a consequence, single-point measurements of powdered samples are disfavored and averaging the signal acquired from a large bed of sample with a rotating sample platform is recommended instead (NIRCal 5.4 Software Manual).

Normalization techniques compensate for scattering by acknowledging that this problem has both additive (baseline-shifting) and multiplicative (intensity dependent) components (Naes et al., 2002). Standard normal variate (SNV) transformation of a spectrum first involves taking the mean and standard deviation of intensities at all wavenumbers in the spectrum. Normalization of each spectrum by SNV is accomplished by subtracting the mean from the intensity at each wavenumber and then dividing by the standard deviation (Naes et al., 2002). The result standardizes the baseline of each spectrum and compensates for intensity fluctuations from scattering, significantly clarifying the basis for comparison of multiple spectra (Figure 1).

2.2.1.2 Smoothing

Smoothing functions are used to remove noise from the signal obtained during scanning of a spectrum (Brereton, 2003). The simplest of these functions is a moving average, in which the intensity measured at a given wavelength is averaged with the measurements obtained at the surrounding wavelengths as shown in Table 1. The intensity of random noise included in any single measurement is dissipated by distributing the error over the window of wavelengths included in the moving average (Brereton, 2003).

Table 1. Windows for smoothing functions.

Moving Average	Smoothed Intensity at λ_k
3-Point	Average (λ_{k-1}, λ_k, λ_{k+1})
5-Point	Average (λ_{k-2}, λ_{k-1}, λ_k, λ_{k+1}, λ_{k+2})
7-Point	Average (λ_{k-3}, λ_{k-2}, λ_{k-1}, λ_k, λ_{k+1}, λ_{k+2}, λ_{k+3})
9-Point	Average (λ_{k-4}, λ_{k-3}, λ_{k-2}, λ_{k-1}, λ_k, λ_{k+1}, λ_{k+2}, λ_{k+3}, λ_{k+4})

Unfortunately, as a linear approximation, the moving average method typically underestimates peak intensity since peaks are better modeled by polynomials (Brereton, 2003). Also problematic is that polynomial calculations are computationally intensive. To address these issues, a simplified calculation to approximate polynomials was determined (Savitzky and Golay, 1964). Savitzky-Golay smoothing calculates a set of coefficients based solely upon the order of the polynomial approximated and the size of the window used for the smoothing operation. The intensity measured at each wavelength in the smoothing window is multiplied by the corresponding coefficient, and the smoothed value is simply the sum of these products.

An example of smoothing a spectrum with error from random noise, represented by an erroneous peak at wavelength 5, is shown in Figure 2. The 5 point moving average function is only slightly worse than Savitzky-Golay functions at minimizing the error at wavelength 5, but much worse at modeling the actual peak centered nearby at wavelength 8. In addition to those for use with original spectra, Savitzky-Golay coefficients have been developed for use with derivative spectra (Brereton, 2003). A number of conceptually similar window-based smoothing operations also exist (Brereton, 2003; NIRCal 5.4 Software manual).

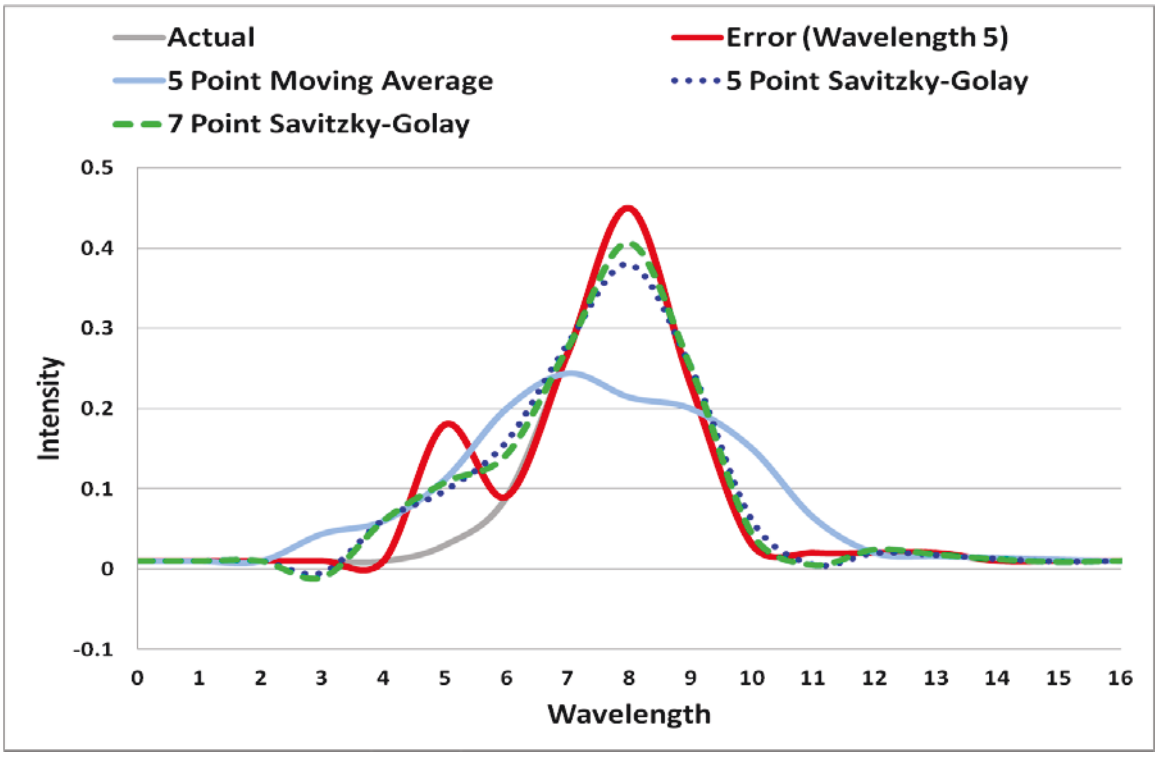


Figure 2. Effect of smoothing on spectra.

2.2.1.3 Derivatization

Derivatization of the original spectrum can enhance features as shown in Figure 3. With each successive derivatization, the peaks of overlapping bands can be separated and intensified. Advantageously, inversion of the second derivative of a signal places its peaks in phase with those of the original spectrum (Naes et al., 2002). However, there are limits as derivatization enhances noise as well as the signal, which can complicate analysis and interpretation (Naes et al., 2002). It is also noteworthy that derivatization inherently adjusts the baselines of spectra to a common value.

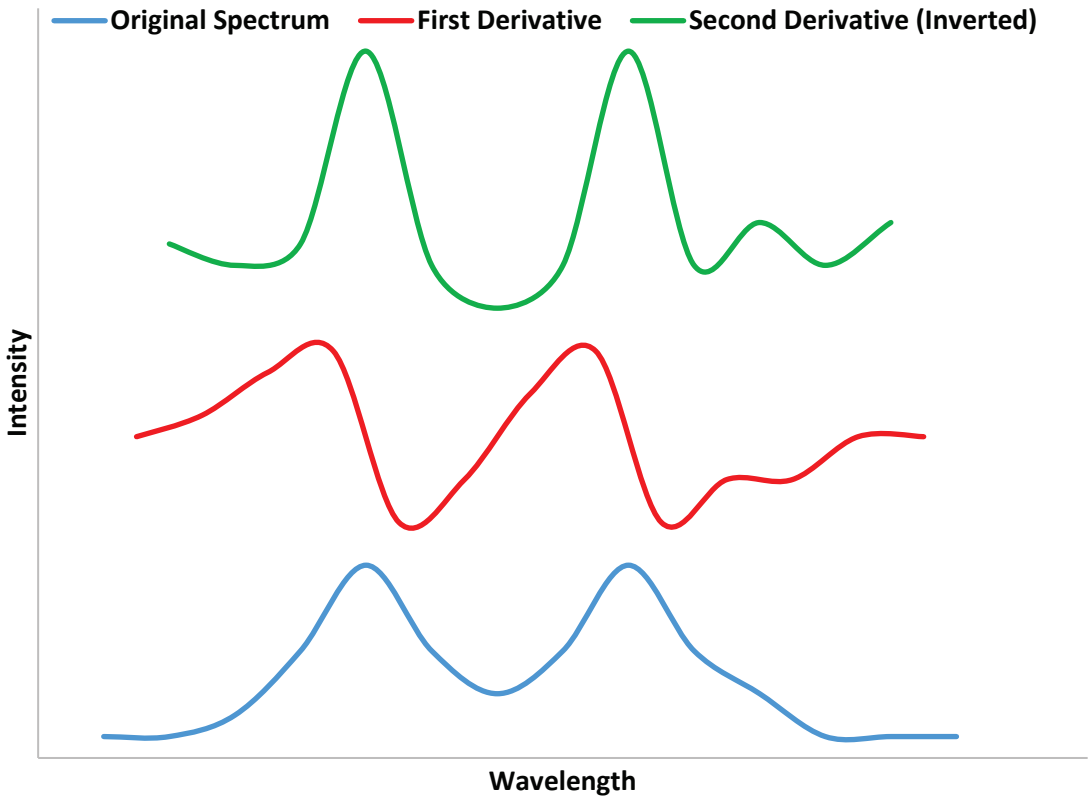


Figure 3. Effect of derivatization on spectra.

2.2.2 Vector Representation of Spectra and Linear Transformation Using Matrices

The main utility of chemometrics is to simplify complex data. In NIR analyses, a spectrum is comprised of a large number of measurements at individual wavelengths. As shown in Figure 4, any spectrum can be represented as a vector in Cartesian space having an axis corresponding to each wavelength (NIRCal 5.4 Software manual). Although this simple example involves only three wavelengths, the Buchi NIRFlex instrument scans 1501 wavelengths, indicating the advantage of chemometric algorithms designed to reduce the dimensionality of this vector representation. Such algorithms operate to rotate and scale the vector representation in space by transforming the original coordinate system into a new one where axes correspond to

linear combinations of the original variables actually measured, known as "latent variables." It is often the case that the dimensionality of the original representation can be reduced, as shown in Figure 4 where the rotated and scaled vector lies along a single dimension.

For simplicity, Figure 4 shows the effect of this transformation on a single spectrum. In practice, multiple vectors corresponding to the spectra in a set of samples are involved as shown in Figure 5. Spectra may be pretreated and measurements at each wavelength are then mean centered prior to transformation. Mean centering results in a vector representation solely indicative of the variation existing among the sample spectra.

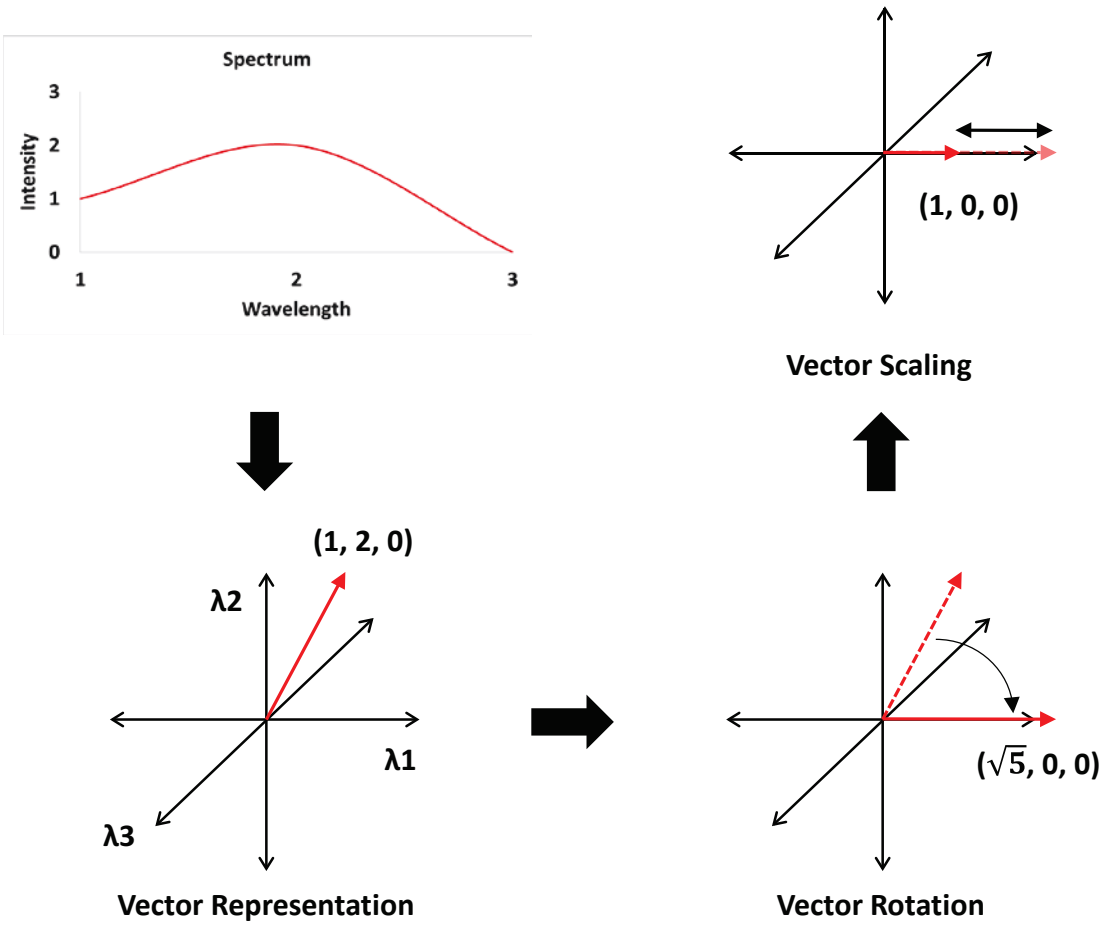


Figure 4. Effect of linear transformation on the vector representation of a spectrum.

In chemometrics, matrices are used to make the linear transformations described above (Naes et al., 2002). For a set of samples, the mean centered intensities measured at a given wavelength represent the variation among the samples along one of the axes in the original coordinate system. These intensities can be collected in a vector $\mathbf{v_1}$. Multiplying $\mathbf{v_1}$ with a matrix **M** transforms or "maps" $\mathbf{v_1}$ into another vector $\mathbf{v_2}$:

$$\mathbf{M}\mathbf{v_1} = \mathbf{v_2}$$

The new vector $\mathbf{v_2}$ represents the variation among the samples along a new axis forming part of the new coordinate system.

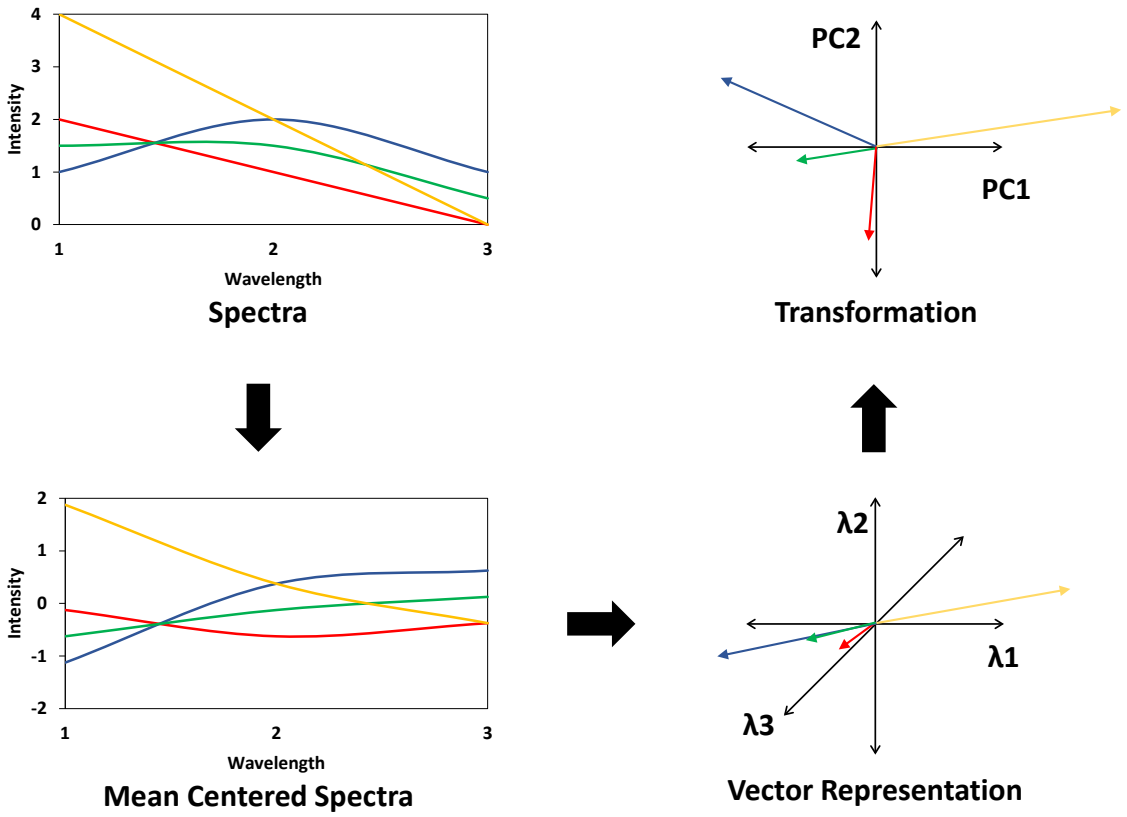


Figure 5. Linear transformation applied to a set of samples.

2.2.3 Results of Linear Transformation: Latent Variable Isolation

Linear transformation generates a new coordinate system in which each vector representing the variation from the mean for a particular sample spectrum in Cartesian space is recast. Each axis in the new coordinate system is a linear combination of the wavelength axes of the original Cartesian space. This linear combination is defined by weight factors indicating the importance of each of the original axes to the new axis. These weights are called loadings (Naes et al., 2002; Brereton, 2003). The placement of each vector in the new coordinate system is given by new coordinates for the corresponding sample called scores (Naes et al., 2002; Brereton, 2003).

2.2.3.1 Loadings

The loadings for each of the two principal components from the linear transformation example of Figure 5 are shown in Figure 6. Loadings for a given principal component are often shown as spectra since there is a value for each wavelength from the original set of x variables (NIRCal 5.4 Software Manual). As different moieties within molecules absorb at characteristic wavelengths, loadings aid in identifying the relationship between each principal component and the chemical composition of the samples. Comparisons are made with the use of a Colthup chart which indicates chemical moieties and their corresponding absorbances in the NIR region. A Colthup chart comprising observations from the literature on the use of NIR in lipid oxidation studies is provided in Appendix A.

2.2.3.2 Scores

The scores for each of the two principal components from the linear transformation example of Figure 5 are shown in Figure 7. This scores plot shows the location of the four samples used in the example in principal component space. The points in the scores plot correspond to the vertices of the corresponding transformed vectors from Figure 5.

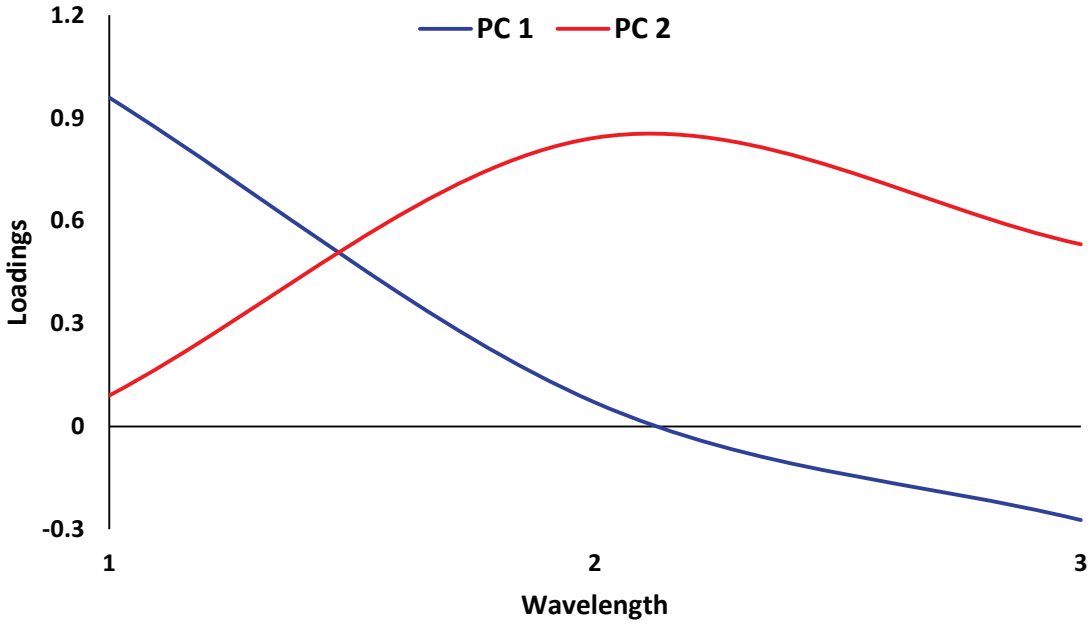


Figure 6. Loadings plot.

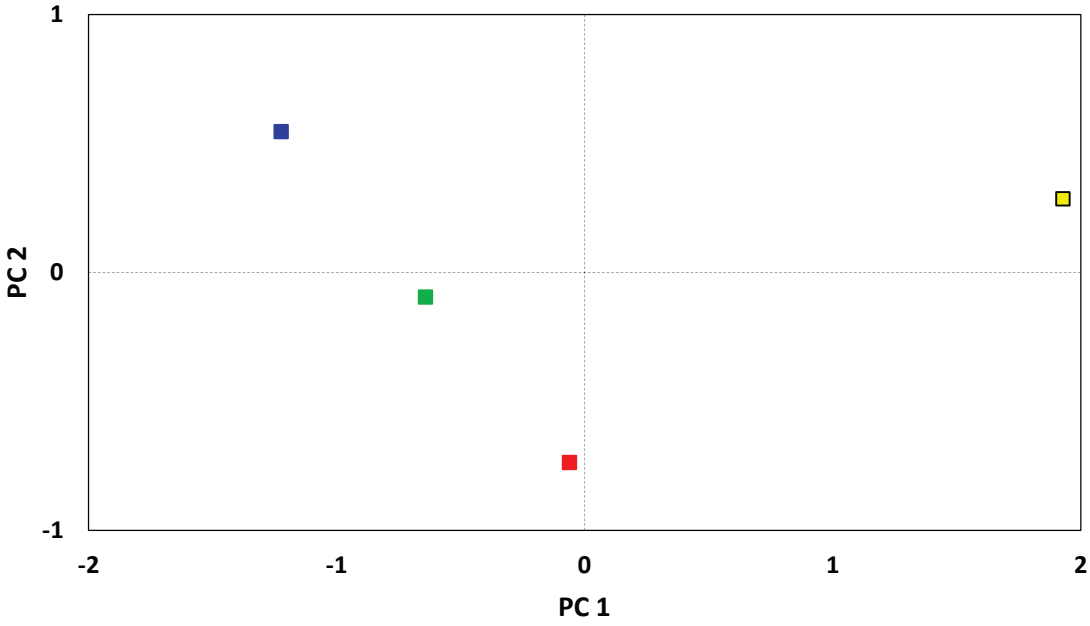


Figure 7. Scores plot.

2.3 Limitations in the Application of NIR to Food Analysis

The examples provided above are both simplified and idealized. In practice, FT-NIR analyses involve far more wavelengths as well as numerous replicate measurements of each sample to ensure statistical significance. Small differences in replicate spectra of a sample can result in variation among the corresponding scores. These differences often arise from random errors inherent in spectroscopic analyses, but can also arise from the nature of the sample.

Depending on the analyte of interest, the variation among scores of nominally identical samples may render analysis impossible. When samples differ greatly in comparison to the variation among replicates, separation of clusters of scores is possible (Figure 8a). When changes among the samples assayed are similar in magnitude to the variation among replicates, analyses can fail as shown in Figure 8b where the samples on the left are unable to be resolved.

Most solid foods are inherently heterogeneous. Thus, spectroscopic assay of different locations in the same solid food sample will result in different spectra due to differences in composition. Additionally, as a physical technique, FT-NIR is affected by light scattering, which increases dramatically for granular and powdered solids. These compositional and physical sources of variation limit the resolving power of FT-NIR analyses of many solid foods. Thus, while FT-NIR has been used for the assay of bulk constituents of solid foods (Osborne & Fearn, 1986), its use for trace constituents has been problematic (see, e.g. Dellarosa et al., 2015).

A common practice used to minimize variation in granular or powdered solid food samples is rotational averaging (NIRCal 5.4 Software Manual). This involves rotating a bed of the sample, often in a petri dish, past the detector as spectra are acquired. The output is a single average spectrum derived from scans of different points of the sample acquired as it is swept past the detector. Although rotational averaging helps to reduce variation it typically requires large amounts of any given sample. This can be prohibitive in a research environment in which many different samples are systematically investigated and only small quantities are available.

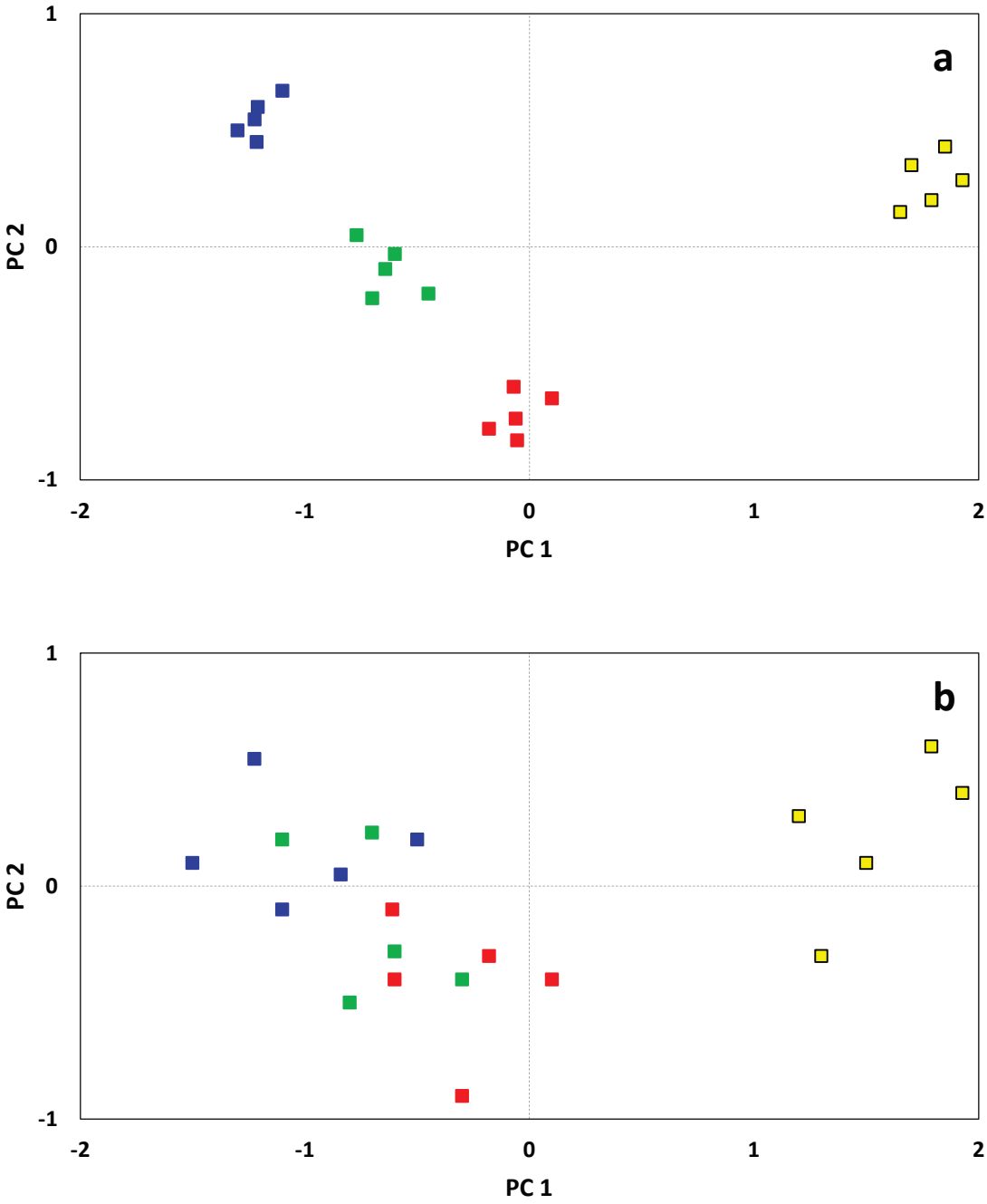


Figure 8. Resolution of samples with different relative variation among replicates: Variation (a) small and (b) large relative to difference between samples.

2.4 Chemometric Approaches to Construct and Assess NIR Models

An overview of the various chemometric techniques used to generate NIR models is presented in Figure 9. Details of the statistical measures used to assess model fitness during validation are shown in Appendix B.

In this investigation, we employed Principal Components Analysis (PCA) for qualitative assessments in pilot studies to discriminate among different lipids incorporated in model solid food systems, among lipids at different stages of oxidation incorporated in model solid food systems, and among different sample forming pressures and modes of presentation (stationary vs. rotating) to the NIR. We employed Partial Least Squares Regression (PLSR) for quantitative assessments of indicators of lipid oxidation in model solid food systems using data from chemical reference assays.

2.5 Review of the Literature: NIR Studies of Lipid Oxidation

2.5.1 Lipid Hydroperoxides

NIR spectroscopy has been used for the determination of peroxide values in both edible oils and food products. A number of research groups have reported varying degrees of correlation between NIR calibrations and reference methods for peroxide values in investigations of neat oils. In solid food systems, correlations are typically hindered by the reduced amount of analyte, interferences from other constituents of the food matrix and the nonhomogeneous nature of foods.

2.5.1.1 NIR Analyses in Neat Oils

The most prevalent application of NIR to determine peroxides in lipids has been in the study neat oils. Mixed results have been reported from studies of different lipids using different oxidation conditions, sample preparation and presentation techniques. Different treatments of spectroscopic data also affected outcomes, with most studies indicating optimal pretreatments.

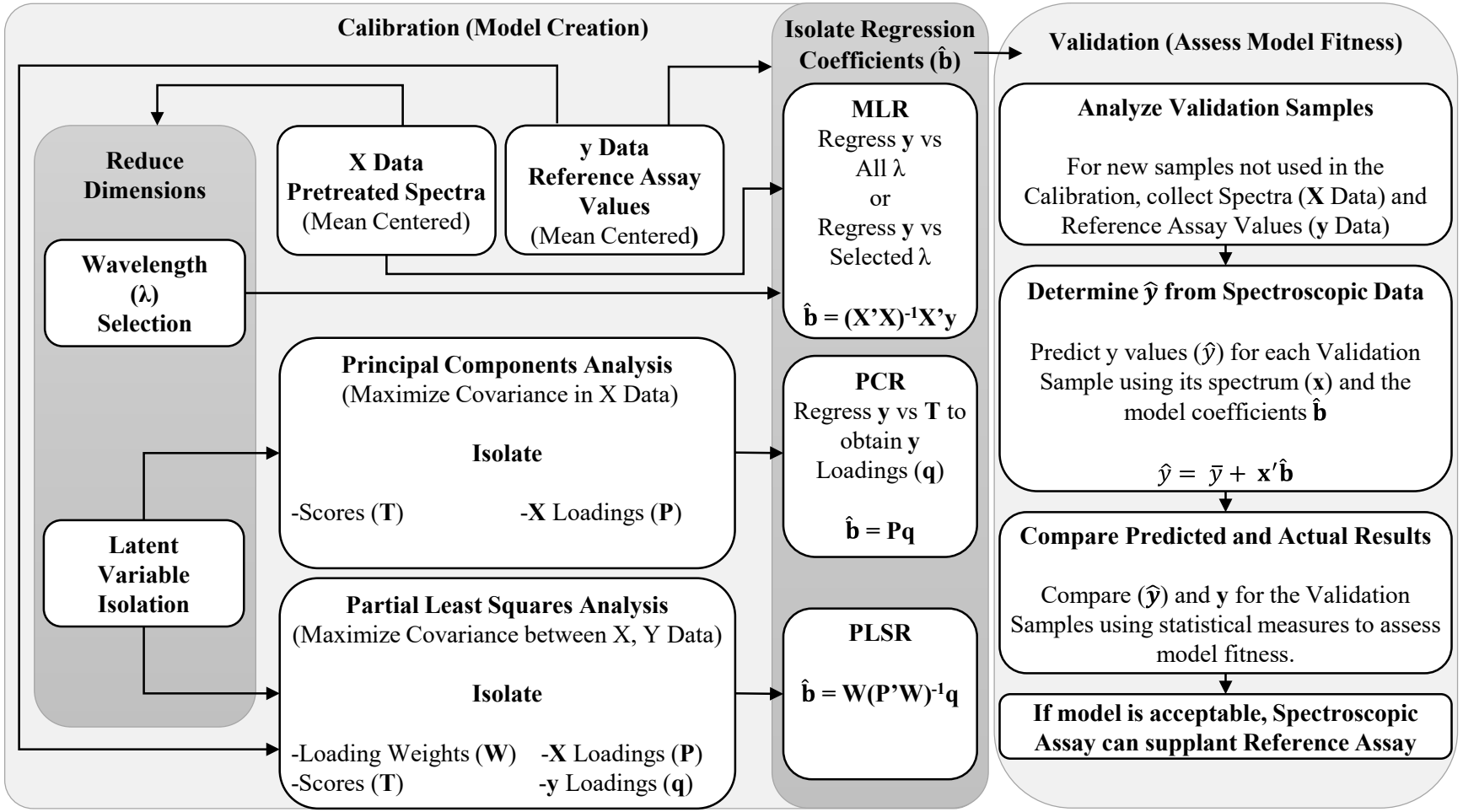


Figure 9. Overview of techniques for chemometric modeling of spectral data.

2.5.1.1.1 Poor to Fair Quality NIR Models

Studies in which NIR models of poor to fair quality for peroxide values were found in neat lipids are shown in Table 2. Each of the four investigations cited used iodometric titration as the reference assay, whether described in EU regulation EEC/2568/91 (Armenta et al., 2007), AOAC standard methods from 1990 (Cozzolino et al., 2005) or 1995 (Szabo et al., 2009), or IUPAC 2.501 (Cayuela Sanchez et al., 2013). Only the study of lard (Szabo et al., 2009) involved oil samples that were thermally treated according to an accelerated oxidation study. Thus, the ranges of peroxide values in many of these studies were smaller than those in the accelerated shelf life studies described below. Moreover, the mean value of peroxides was typically skewed toward the low end of each range, further indicating the effect that limited oxidation could have had on these studies.

Each of these studies used partial least squares to model pretreated data from large spectral ranges. Armenta et al. (2007) used 4599-5450 cm^{-1} and 7498-11959 cm^{-1} for both olive oil and sunflower, seed and maize oil models and 4550-5450 cm^{-1} and 6100-7500 cm^{-1} for the combined model, each of which relied upon first derivative and normalization pretreatments. Cozzolino et al. (2005) used 4000-9091 cm^{-1} (1100-2500 nm) with second derivative treatment and Savitzky-Golay smoothing. Szabo et al. (2009) used 4000-12,500 cm^{-1} with multiplicative scatter correction, second derivative and smoothing pretreatments. Cayuela Sanchez et al. (2013) applied variable selection techniques to choose wavenumber ranges from within an overall range of 4000-28,571 cm^{-1} with normalization of absorbance data as the only pretreatment. That group appeared to have the best validation results based solely upon the RPD value.

Interestingly, two studies (Armenta et al., 2007; Cayuela Sanchez et al., 2013) used transmittance instruments with increased path lengths than those used in the transflectance instruments of the others (Cozzolino et al., 2005; Szabo et al., 2009). Thus, both modes of sample presentation were applied in this group. The number of samples used also ranged from

moderate (33 in Armenta et al., 2007 and Szabo et al., 2009) to large (245 in Cayuela Sanchez et al., 2013). Other factors such as extent of oxidation must have adversely affected these results.

It was noteworthy that all but the Armenta et al. (2007) study exhibited fair to poor linearity among their calibration data sets (r^2 in Table 2), which typically indicates assay condition problems. Cozzolino et al. (2005) stated that operational constraints caused at least a 5 day delay between the NIR and reference analyses, during which time samples were stored in the dark at room temperature. Changes in peroxide values between assays were thus the likely reason for the inability to obtain a linear calibration. Szabo et al. (2009) found a calibration r^2 of only 0.483 for samples subjected to stability testing. However, that study applied eight frying temperatures (160-230°C) for eight hours per day for up to four days. The authors attributed the weak correlation to the rapid decomposition of peroxides under their assay conditions. A dramatic increase in peroxide values was observed after the first eight hours of incubation which was then followed by a drop almost to initial levels.

Interestingly, Cayuela Sanchez et al. (2013) reported a calibration r^2 of only 0.87 using 199 calibration samples, quartz apparatus for sample presentation and computerized variable selection techniques to optimize wavenumber ranges used in the model. A box and whiskers plot of the calibration samples indicated approximate values of Q1, Q2 and Q3 were 9, 12, and 26 meq/kg. Thus, although the overall range of peroxide values was moderately large (up to 43.0 meq/kg), the bulk of the samples were in the less oxidized area of the range. The authors noted that the high monounsaturated to polyunsaturated fatty acid ratio and the natural presence of strong antioxidants in virgin olive oil afforded high resistance to oxidative degradation. The fair linearity among calibration samples may indicate issues regarding sensitivity of the technique and the model may be improved by incorporating samples with higher levels of oxidation.

Table 2. Poor to fair quality NIR models for peroxide values in lipids from the literature.

Reference	Lipid	Number of Samples	Sample Presentation	Range (meq / Kg)	Model (Factors)	RMSEC (meq / Kg)	RMSEP (meq / Kg)	SEC (meq / Kg)	SEP (meq / Kg)	Calibration r^2	Validation r^2	RPD
Armenta et al., 2007	Olive Oil	33 (14 Calibration 10 Validation 9 Prediction)	Transmittance in 6.5 mm i.d. Glass Vials in 30 C Thermostatized Cell Holder	2.5 - 17.2 Validation (Mean 8 SD 3) Prediction (Mean 6 SD 2)	PLS (7)	0.33	1.87 (1.20) (RMSEV)			0.997	0.656	1.1 (2.5) (Val)
	Sunflower, Seed and Maize Oils	36 (15 Calibration 12 Validation 9 Prediction)		1.8 - 7 Validation (Mean 3.5 SD 0.5) Prediction (Mean 3.5 SD 0.8)	PLS (7)	0.23	0.79 (0.63) (RMSEV)			0.980	0.500	1.0 (0.8) (Val)
	Combined	69 (29 Calibration 22 Validation 18 Prediction)		1.8 - 17.2 Validation (Mean 5 SD 2) Prediction (Mean 5 SD 3)	PLS (10)	0.94	1.50 (1.20) (RMSEV)			0.950	Not Reported	2 (1.7) (Val)
Cozzolino et al., 2005	Oil Byproduct of Fishmeal	160 (80 Calibration 80 Validation)	Transflectance in 0.1 mm Aluminum Back Quartz Camlock Cell	0.12 - 16.0 (Mean 5.2 SD 2.8)	PLS (Not Reported)				3.9	0.6	0.16	0.7
Szabo et al., 2009	Lard	33 ASL Treatments of 1 Sample	Transflectance in 0.1 mm Aluminum Back Quartz Camlock Cell	3.56 - 130.0 (Mean 30.9 SD 31.3)	PLS (1)			22.51	24.05 (SECV)	0.483	Not Reported	1.3
Cayuela-Sanchez et al., 2013	Virgin Olive Oil	245 (199 Calibration 46 Validation)	Transmittance in 5 mm Path Length Quartz Cuvettes	5.6 - 43.9 Calibration (Mean 17.0 SD 10.7) Validation (Mean 18.7 SD 10.8)	PLS (Not Reported)			3.81	3.82	0.87	Not Reported	2.84

2.5.1.1.2 Improved Quality NIR Models from ASL Studies

Studies that employed accelerated shelf life (ASL) testing generally increased the range of sample oxidation and resulted in improved models for peroxide values in neat oils. These studies used various thermal or illumination conditions to oxidize samples of olive, sunflower, rapeseed, soybean, canola, safflower or cottonseed oils. More recent studies applied partial least squares, while previous ones found single or double wavelength models to be optimal after application of multiple linear regression to spectral data.

Studies that found better PLS models for peroxide values in neat oils are summarized in Table 3. Each used iodometric titration as the reference assay, whether by AOCS Official Methods Cd 8b-90 (Manley & Eberle, 2006) or Cd 8-53 (Yildiz et al., 2001), or by EU Regulation 61/2011, L23 (Wojcicki et al., 2015). All were ASL studies. Manley and Eberle (2006) used moderately elevated temperatures with illumination for up to 10 weeks on extra virgin olive oil. Wojcicki et al. (2015) applied higher temperatures in darkness for up to only 15 days on olive, sunflower and rapeseed oils. Yildiz et al. (2001) used illumination with fluorescent light for up to 180 hours (7.5 days) to oxidize soybean oil. Peroxide value ranges varied with ASL conditions and oil type, but were generally larger than those of studies cited in Table 2.

PLS models were successfully generated using both transmittance and transflectance modes for sample presentation. Manley and Eberle (2006) compared results using two different FT-NIR spectrometers, a Perkin Elmer IdentiCheck and a Buchi NIRLab N-200. Measurements on the Perkin Elmer spectrometer were made at four different resolutions (64, 32, 16 and 8 cm^{-1}) using quartz cuvettes at each of 2 path lengths (0.2 mm and 0.5 mm) in transmittance mode. Measurements on the Buchi spectrometer were performed in transflectance mode at a higher resolution (4 cm^{-1}) using glass Petri dishes fitted with a 0.3 mm high reflectance plate affording a path length of 0.6 mm overall. Although paired comparison tests indicated no significant differences among peroxide value model quality using different path lengths and scanning resolutions, significant differences were observed among the two spectrometers respecting the

standard deviation of repeated measurements. The instrument which relied upon quartz cuvettes resulted in reduced standard deviation among measurements which was attributed to the consistent path length provided. Measurements from the instrument which relied upon Petri dishes for sample presentation indicated fluctuations arising from different path lengths due to their irregular surfaces. The authors noted that transmittance would be preferable to transflectance in cases such as this where the sample presentation apparatus affected variation among measurements. Sample presentation in both of the other PLS model studies was via transmittance using longer path lengths. Yildiz et al. (2001) relied upon both 1 and 2mm quartz cuvettes, with better validation results arising from the latter. Wojcicki et al. (2015) described their measurement cell as an 8 mm wide glass vial.

Optimized models in extra virgin olive oil (Manley & Eberle, 2006) were obtained using only normalization as a pretreatment from 4000-9091 cm^{-1} (1100-2500 nm) for the transmittance instrument and 4000-10,224 cm^{-1} (978-2500 nm) for the transflectance instrument. Absorbance increases throughout the NIR spectra were observed with oxidation, with particular effect between 4545 and 5000 cm^{-1} attributed to the formation of unsaturated hydroperoxides.

Although Wojcicki et al. (2015) investigated the use of both mid and near IR, the best models for peroxides were found to arise solely from the NIR region. Peroxide models in olive, sunflower and rapeseed oils used mean-centered absorbance spectra from 4000-12,500 cm^{-1}. PLS models for sunflower and rapeseed oil required fewer factors than the model for olive oil. Peroxide formation was fastest in sunflower oil, followed by refined rapeseed oil, cold-pressed rapeseed oil, and olive oil in keeping with fatty acid content and levels of natural antioxidants. As samples oxidized, spectral absorbances rose at 4810 cm^{-1} as well as on the lower energy side of a peak at 7068 cm^{-1}, which correlated with characteristic hydroperoxide bands at 4831 cm^{-1} and 6849 cm^{-1} determined by Holman (Holman et al., 1958). Principal component analysis on a combined sample set of all oils indicated that 90% of spectral variation was accounted for by two principal components. Loadings of PC1, which accounted for 75% of the variation, included

bands at 4800 cm^{-1} and 6852 cm^{-1}, both attributable to the generation of hydroperoxides. PC2 appeared to correlate with unsaturated structures lost upon oxidation.

The optimized PLS model for peroxides in soybean oil (Yildiz et al., 2001) used first derivative pretreatment of the range from 1100-2200 nm (4545-9091 cm^{-1}). Key changes among oxidized samples were observed at 1200-1400 nm (7143-8333 cm^{-1}), 1700 nm (5882 cm^{-1}), and 2000-2200 nm (4545-5000 cm^{-1}). These features were attributed to OH stretching vibrations, CH stretching vibrations, and OH combination bands, respectively. Although the range of peroxide values was almost as narrow as studies from Table 2, validation statistics were greatly improved likely due to the more even distribution of peroxide values in this ASL study. The number of PLS factors (13) was highest among the studies in Table 3 given the diminished extent of oxidation, in keeping with the observation of Wojcicki et al. (2015).

Yildiz et al. (2003) subsequently used this PLS model in a comparison among iodometric titration (AOCS Official Method Cd 8-53) with NIR as well as the ferrous xylenol orange (FOX), and PeroxySafe™ chemical assays in five soybean oils oxidized with fluorescent light for up to 216 hours. Hydroperoxides determined by iodometric titration correlated highly with both their NIR method ($r = 0.991$ and the Standard Deviation of Differences (SDD) = 0.72 meq/kg) and the PeroxySafe™ method ($r = 0.993$ and SDD = 0.56 meq/kg). They observed a weaker correlation between iodometric titration and the FOX method ($r = 0.975$ and SDD = 2.3 meq/kg). Accordingly, they concluded that iodometric titration, NIR and the PeroxySafe™ assay were equivalent for the determination of peroxide values in soybean oil

Table 3. Partial Least Squares (PLS) NIR models for peroxide values in lipids from the literature.

Reference	Lipid	Number of Samples	ASL Conditions	Sample Presentation	Range (meq / Kg)	Model (Factors)	RMSEP (meq / Kg)	SEP (meq / Kg)	Calibration r^2	Validation r^2	RPD
Manley & Eberle, 2006	Extra Virgin Olive Oil	90 Calibration and 44 Validation from Treatment of 23 Oils	35°C With Illumination for up to 10 Weeks	Transmittance in 0.2 mm Quartz Cuvettes	Calibration 2.18 - 74.02 (Mean 20.51 SD 15.40) Validation 2.47 - 60.62 (Mean 20.16 SD 14.62)	PLS (6)	4.16	4.15		0.92	3.52
	Extra Virgin Olive Oil			Transflectance using glass Petri dishes with 0.3 mm high reflector		PLS (8)	5.29	5.28		0.87	2.77
Wojcicki et al., 2015	Olive Oil	36 (6 Treatments of 6 Oils)	60°C in Darkness for up to 15 Days	Transmittance in 8 mm path length glass vial	5.9 - 53.2 (Mean 27 SD 17)	PLS (6)	2.5 (RMSECV)			0.977	6.8
	Sunflower Oil	12 (6 Treatments of 2 Oils)			19 - 834 (Mean 335 SD 298)	PLS (2)	42.4 (RMSECV)			0.982	6.3
	Rapeseed Oil	30 (6 Treatments of 5 Oils)			1.1 - 568 (Mean 163 SD 166)	PLS (3)	39.3 (RMSECV)			0.945	4.2
	Combined	78 (6 Treatments of 13 Oils)			1.1 - 834 (Mean 127 SD 188)	PLS (3)	32.1 (RMSECV)			0.970	5.3
Yildiz et al., 2001	Soybean Oil	128 (16 Treatments of 8 Oils) (75 Calibration and 44 Validation)	Fluorescent Light (12 hr Intervals up to 180 hrs)	Transmittance in 2 mm path length Quartz Cuvettes	Approximately 0.2 - 23 Calibration (Mean 9.60 SD 6.18) Validation (Mean 10.23 SD 6.32)	PLS (13)		0.720 (0.760 for External Validation Set)	0.992	0.988 (0.994 for External Validation Set)	8.8

Table 4. Single and double wavelength NIR models from MLR for peroxide values in lipids from the literature.

Reference	Lipid	Number of Samples	ASL Conditions	Sample Presentation	Range (meq / Kg)	λs from MLR	SEP (meq / Kg)	Calibration r^2	Validation r^2	RPD
Cho et al., 1998	Soybean Oil	35 Calibration and 15 Validation	75°C for up to 3 Weeks	NIRSystems Model 4500	Calibration 0.4 - 376.3 (Mean 139.84) (SD 112.94) Validation 0.7 - 308.1 (Mean 124.94) (SD 111.09)	(2 λs) 2020 nm and 2080 nm 4950 cm^{-1} and 4808 cm^{-1}	9.67	0.994	0.992	11.5
Takamura et al., 1995	Canola Oil	14 - 15 Validation Samples for Each Assay	50°C until reaching PV of 600 meq / Kg	Transmittance in 1 mm path length cuvette cell Thermostatized at 30°C	Approximately 0 to 600	(1 λ) 2084 nm 4798 cm^{-1}	8.5		0.998	
	Olive Oil						12.7		0.992	
	Safflower Oil						19.2		0.990	
	Canola Oil					(1 λ) 1744 nm 5734 cm^{-1}	18.4		0.988	
	Olive Oil					(1 λ) 1746 nm 5727 cm^{-1}	19.5		0.982	
	Safflower Oil					(1 λ) 1742 nm 5741 cm^{-1}	24.0		0.986	
	Soybean Oil	6 - 7 Validation Samples for Each Assay	50°C until reaching PV of 30 meq / Kg		Approximately 0 to 30	(1 λ) 2084 nm 4798 cm^{-1}	0.83		0.992	
	Cottonseed Oil						1.37		0.988	

Studies that found good models for peroxide values in neat oils using one or two wavelengths determined from multiple linear regression (MLR) are summarized in Table 4. In a three week accelerated shelf life study of soybean oil held at 75 °C, Cho et al. (1998) found a model incorporating two wavelengths, primarily 2080 nm (4808 cm^{-1}) and secondarily 2020 nm (4950 cm^{-1}), was not significantly improved upon by the inclusion of additional wavelengths. Second derivative pretreatment of spectra was used to separate the signals arising from other chemical changes during oxidation from those of interest. Validation statistics indicated excellent fit ($r^2 = 0.992$ and RPD = 11.5). This model was highly accurate in the early stages of lipid oxidation during peroxide accumulation; however, during longer periods where peroxide decomposition occurred, the accuracy of the model dropped. This phenomenon was previously observed (Hong et al., 1994) and was later cited by Szabo et al. (2009) to explain the poor performance of NIR in assessing peroxides in oils subjected to frying temperatures long beyond the point at which they decomposed.

An accelerated shelf life study by Takamura et al. (1995) followed peroxide values up to 600 meq/kg in stripped canola, olive and safflower oils and up to 30 meq/kg in soybean and cottonseed oils subjected to autooxidation at 50 °C. Peak intensity at 2084 nm (4798 cm^{-1}) in second derivative spectra correlated highly with peroxide values in each of these oils. Models using only this wavelength for each oil exhibited $r^2 \geq 0.988$ and small standard errors of prediction for the ranges of peroxides assessed (8.5, 12.7 and 19.2 meq/kg for canola, olive and safflower oils and 0.83 and 1.37 meq/kg for soybean and cottonseed oils). The authors noted that Holman et al. (Holman et al., 1958) reported lipid peroxides at 1460 and 2070 nm (6849 and 4831 cm^{-1}), and that the difference between the latter and the 2084 nm (4798 cm^{-1}) peak they observed was likely due to the pretreatment applied. They also confirmed that this peak was due to hydroperoxides using pure systems of methyl linoleate autooxidized at 50 °C and observing losses in this peak in methyl linoleate hydroperoxide upon reduction with sodium borohydride.

Among the more easily oxidized oils, single wavelength models from another area of the spectrum also correlated well with peroxide values (1746 nm (5727 cm^{-1}) in olive oil, 1744 nm (5734 cm^{-1}) in canola oil, and 1742 nm (5741 cm^{-1}) in safflower oil).

The importance of the band in the 4800 cm^{-1} range determined by both the Cho and Takamura groups as well as that in the 5700 cm^{-1} range determined by Takamura was reinforced by Yildiz et al. (2001). In their soybean oil study discussed above, the Yildiz group also performed multiple linear regression and found the best four wavelength models at both 1 mm and 2 mm path lengths. The 1 mm model included 2070 nm (4831 cm^{-1}), 2036 nm (4912 cm^{-1}), 1746 nm (5727 cm^{-1}), and 1400 nm (7143 cm^{-1}), while the 2 mm model included 2068 nm (4836 cm^{-1}), 2016 nm (4960 cm^{-1}), 1612 nm (6203 cm^{-1}), and 1242 nm (8052 cm^{-1}). In both cases, the band around 2070 nm (4831 cm^{-1}) was highly significant, while the band at 1746 nm (5727 cm^{-1}) appeared in the model for the same 1 mm path length used in the study by the Takamura group. Despite agreement with the Takamura group's finding that 2084 nm was critical, Yildiz et al. (2001) cautioned against a single wavelength approach given the potential of interference from a variety of other hydroxyl containing compounds.

2.5.1.2 NIR Analyses in Complex Food Systems

Attempts to monitor peroxide values in solid food systems via NIR have generally been unsuccessful. The food matrix complicates spectra by reducing signal intensity for analytes of interest relative to their levels in neat oils, adding potential interferences from signals of matrix constituents and giving rise to homogeneity issues resulting in increased variation among measurements of the same sample in different areas. The matrix also complicates correlation between NIR and chemical reference assays as the latter typically require lipid extraction prior to analysis, a process which can lead to differences between the NIR sample and the reference assay sample. These problems require the application of controls over sample presentation during scanning as well as consideration of what constitutes a representative sample and what spectral regions should be used to construct models.

The studies summarized in Table 5 are generally indicative of these complications. Jensen et al. (2001) investigated walnut kernels and Dellarosa et al. (2015) investigated seven formulations of precooked fish cakes using ferric thiocyanate reference methods for peroxides. Kaddour et al. (2006) investigated three food products, salted crackers, moist Asian noodles, and healthy crackers, as well as bulk rapeseed oil using iodometric titration (AOCS Official Method Cd 23-93) as the reference method.

Although all three were shelf life studies, only Kaddour et al. (2006) was a true ASL, with products subjected to 40 °C for up to two months. The others used chilled (0 to 5 °C) (Jensen et al., 2001; Dellarosa et al., 2015) or room temperature (Jensen et al., 2001) conditions to track oxidation during normal shelf life conditions for those products. Peroxide value ranges in the lower temperature studies were small to moderate, resembling those in studies of neat oils which yielded poor models (Table 2) and likewise implicating the sensitivity of NIR for these analytes. Ranges of peroxide values for the products in the Kaddour et al. (2006) study varied according to fatty acid content (saturated vs. unsaturated) and disposition of lipid within the product (homogeneous throughout or on the surface), and for all products except salted crackers were far higher than those of the other studies in Table 5. NIR models of the most oxidized product investigated, moist Asian noodles, relied on data from only the first half of the study because after a few weeks drying of the product altered spectra and degraded model quality. (Kaddour et al., 2006) Nevertheless, these models had the best validation statistics (RPD) among those in Table 5, likely due to the extent of oxidation.

Sample sizes were moderate, with the models of the Kaddour et al. (2006) ASL using roughly half (17-24) the number as the normal shelf life studies (50 for Jensen et al. (2001) and 45 for Dellarosa et al. (2015)). Jensen et al. (2001) drew duplicate walnut samples at T0 as well as 3, 8, 12, 16 20 and 25 weeks under each of the four assay conditions for chemical analysis. Kaddour et al. (2006) drew samples ‘periodically’ up to 60 days. Dellarosa et al. (2015) drew fish cake samples at T0 as well as 4, 7, 14 and 28 days of storage, the stated shelf life of the product.

Under three of the four conditions studied (storage at 21°C with or without illumination and 5 °C with illumination) by Jensen et al. (2001), peroxide values in walnuts were observed to rise and then fall within the duration of the study. The precooked fish cakes of Dellarosa et al. (2015) were fried and baked prior to the study; peroxides formed during cooking and measured in the T0 samples dropped by the second time point and then rose again with the third. The inclusion of samples from after the point at which peroxides degraded, according to the findings of Cho et al. (1998) and Hong et al. (1994), may also have adversely affected models in both studies.

Due to the volume of work, chemical and NIR assays of walnuts were unable to be performed on the same day (Jensen et al., 2001). Samples were stored at 5 °C in vacuum packed bags in the dark in the intervening time, the duration of which was not specified. Although these conditions were improved over those of Cozzolino et al. (2005) (five days in the dark at room temperature), they were similar to those of Dellarosa et al. (2015) who observed a drop in levels of peroxides in fish cakes during the first four days of chilled storage. Thus, changes in peroxide content of the samples between assays may have impacted the correlation between chemical testing and NIR in models of the walnut study.

Sample presentation in the solid food systems studies occurred by reflectance spectroscopy. Ground walnut kernels were scanned in reflectance mode using a rotating sample cup with a quartz window and a compressive paper disk to ensure constant pressure on the sample (Jensen et al., 2001). Kaddour et al. (2006) scanned products both intact and after grinding using a fiber optic probe. Dellarosa et al. (2015) did not specify the exact mode of sample presentation other than that a Bruker Optik MPA spectrometer was used on minced fish cakes. On the issue of sample homogeneity, scans of walnuts employed rotational averaging, while five scans of distinct areas in stationary samples were averaged for spectra of minced fish cakes. It is questionable whether five stationary points (Dellarosa et al., 2015) or the use of a fiber optic probe (Kaddour et

al., 2006) would result in a representative sample given the heterogeneous nature of foods. The improvement in the fit of calibration observed by Kaddour et al. (2006) for healthy crackers upon grinding (r^2 of 0.711 versus 0.575 for intact samples) is further evidence that sample heterogeneity is an important consideration.

It is also important to consider the spectral regions to use in a model. Excellent correlations for peroxide values have been found in neat oils using single and double wavelengths in the 4800 - 5000 cm^{-1} area as well as the 5730 cm^{-1} area (Cho et al., 1998; Takamura et al., 1995). Jensen et al. (2001) found NIR to correlate poorly with peroxides (validation $r^2 = 0.55$) using second derivative spectra in the combined visible and NIR ranges (400 - 2498 nm or 4003 - 25,000 cm^{-1}). Results using an NIR specific range (1850 - 1980 nm or 5051 - 5405 cm^{-1}) proved much worse (validation $r^2 = 0.28$); however, this range did not include any of the wavelengths indicative of peroxides from other literature and was chosen because it provided the best separation among oxidized walnuts generally in principal component analysis scores plots. This simply indicates that changes in peroxides were not the principal driver to alter spectra of walnuts under the oxidation conditions used in that study.

Full spectrum models were also used in Kaddour et al. (2006) (1000 - 2500 nm or 4000 - 10,000 cm^{-1}) and Dellarosa et al. (2015) (800 - 2500 nm or 4000 - 12,500 cm^{-1}), and the latter also used a slightly truncated region (1100 - 2200 nm or 4545 - 9091 cm-1) based upon the findings of Yildiz et al. (2001). Kaddour et al. (2006) reasoned that many of their models suffered because sample sets were too small. Dellarosa et al. (2015) cited the minimal range of peroxide values and interference from the food matrix for the lack of fit. Although these factors contributed to their results, and in addition to the other issues with these studies noted above, the finding that a small number of wavelengths correlated well with peroxide values (Cho et al., 1998; Takamura et al., 1995) could indicate that the wavelength ranges used to construct these models were too large.

Studies in solid food systems highlight two problems associated with NIR analysis of lipid oxidation products. One is the sensitivity and limit of detection, since as shown in several of the studies cited above, NIR models tend to be poorer when peroxide levels were low or when high levels of other products were also present. Whether this arose directly due to low absorptions from hydroperoxides or indirectly due to interferences from other oxidation products remains to be determined. Another problem inherent in correlating NIR data with chemical analyses is that the latter require extraction of lipids from the solid matrix. This usually alters the endogenous oxidation status (e.g. increasing some products and decomposing others) and, depending on the extraction method, is often incomplete. Thus, analyses of extracts of solid materials may not detect the same products as those determined by NIR of the intact materials.

2.5.2 Conjugated Dienes

NIR spectroscopy has been used to develop models of varying quality for conjugated dienes in soybean oil and olive oils. As shown in Table 6, models were created using PLS as well as combinations of four wavelengths isolated from forward stepwise MLR. These studies primarily relied upon transmittance spectroscopy with quartz optics and used large sample sets.

During the ASL study of soybean oil cited above in the discussion of peroxide values, Yildiz et al. (2001) also assessed NIR models to quantify conjugated dienes. Reference assay values were obtained as percent conjugated dienes by measurement at 233 nm according to AOCS Official Method Ti 1a-64 (1990). As oxidation via fluorescent light proceeded, changes in levels of conjugated dienes tracked those of peroxides. Initial conjugated diene values for the eight soybean oils ranged from 0.20-0.26%, and although a range of conjugated dienes was not reported, the mean of both calibration and validation samples was 0.31% with a standard deviation around 0.1%.

Table 5. Partial Least Squares (PLS) NIR models for peroxide values in solid food systems from the literature.

Reference	Food System	Number of Samples	ASL Conditions	Sample Presentation	Range (meq / Kg)	Model (Factors)	RMSEP (meq / Kg)	Calibration r^2	Validation r^2	RPD
Jensen et al., 2001	Walnut Kernels (53% Fat) Linoleic Acid (62.0%) Oleic Acid (17.9%) Linolenic Acid (10.7%) Palmitic Acid (7.1%) Stearic Acid (2.3%)	50 (2 Samples Each at T0 and 6 Time Points for Each of 4 Conditions)	Storage in Light or Darkness at 5°C or 21°C for up to 25 Weeks	Reflectance in Rotating Sample Cup with Quartz Window and Compressive Paper Disk	0.9 - 12.9 (Mean 5.2 SD 2.9)	PLS (3) (Visible and NIR)	1.9 (RMSECV)		0.55	0.66
						PLS (3) (Visible)	1.5 (RMSECV)		0.68	0.52
						PLS (1) (NIR)	2.5 (RMSECV)		0.28	0.86
Kaddour et al., 2006	Salted Crackers (20% Palm Oil and Lauric Oil on Surface of Product)	23	Storage at 40°C for up to 2 Months	Fiber Optic Probe on Intact Product Samples as well as After Grinding	0.7 - 5.3 (Mean 1.8)	PLS (7 Intact) (3 Ground)		0.726 (Intact) 0.446 (Ground)		1.14 (Intact) 1.03 (Ground)
	Moist Asian Noodles (1.5% Rapeseed Oil on Surface of Product)	17			10.4 - 87.6 (Mean 54.4)	PLS (5 Intact) (7 Ground)		0.897 (Intact) 0.916 (Ground)		1.42 (Intact) 1.40 (Ground)
	Healthy Crackers (10% Rapeseed Oil throughout Product)	24			1.2 - 46.9 (Mean 14.0)	PLS (5 Intact) (5 Ground)		0.575 (Intact) 0.711 (Ground)		1.16 (Intact) 1.20 (Ground)
	Rapeseed Oil	23		Fiber Optic Probe	0.5 - 161 (Mean 43.7)	PLS (4)		0.381		1.03
Dellarosa et al., 2015	Fish Cakes (7 Different Formulations Cooked and Vacuum Packed)	33 Calibration and 12 Validation	Storage at 0-4°C for up to 28 Days (Normal Shelf Life)	Minced Samples Assayed in Bruker Optik MPA Spectrometer	Approximately 0.8 - 3.6 (Dependent upon Formulation)	PLS (9)	0.408		0.738	

Models were made using samples scanned in 1 mm or 2 mm path length quartz cuvettes, with the latter observed to provide improved validation statistics. All soybean oil models reported in Table 6 used first derivative pretreatment of spectra. A discrepancy was observed when wavelength ranges were optimized to provide the best PLS models for each path length. Although the best 2 mm model used 1100 - 2200 nm (4545 - 9091 cm^{-1}), 1 mm models were not improved by excluding the 2200 - 2500 nm range (4000 - 4545 cm^{-1}). Yildiz et al. (2001) proposed that nonlinearities observed at longer path lengths in that region were likely responsible. As a result, the wavelength range for the best 1 mm path length model of conjugated dienes differed from that shared by the best 2 mm path length model and the best models of peroxides at both path lengths.

In addition to indicating the importance of a primary wavelength in the 2064-2070 nm region (4831-4845cm^{-1}), four wavelength MLR models for conjugated dienes using data from each path length indicated discrepancies based on sample thickness (Table 6). Two of the four wavelengths for the 1 mm model, 2430 nm (4115 cm^{-1}) and 2350 nm (4255 cm^{-1}), fell within the region excluded by PLS models of the thicker sample. The remaining two wavelengths were shared, including the primary wavelength and one at 1396-1398 nm (7153-7163 cm^{-1}).

Validation statistics indicated PLS and MLR models were fair for 1 mm and good for 2 mm path length. They were also robust as shown by good results obtained with an external validation set of thirty samples made from ASL treatment of three additional oils (Table 6). Although good correlation of NIR with conjugated dienes was observed, Yildiz et al. (2001) noted that it was not as strong as that obtained for peroxides (see Table 3).

Manley and Eberle (2006) investigated conjugated dienes in extra virgin olive oil samples oxidized for up to 10 weeks at 35°C using the specific extinction coefficient at 232 nm (K232) (Table 6). As they observed for peroxides, paired comparison tests showed a statistically significant reduction in the standard deviation of measurements of conjugated dienes from NIR spectra acquired via transmittance in quartz cuvettes relative to those by transflectance in glass

petri dishes. The best PLS model based upon transmittance relied upon the 4000-9091 cm^{-1} range, had only six factors and provided a validation r^2 of 0.94 and a RPD of 3.62 indicating it was suitable for screening. The best PLS model based upon transflectance used the 4000-10,224 cm^{-1} range, had only four factors and provided a validation r^2 of 0.87 and a RPD of 2.56 indicating it was only suitable for very rough screening. They did not elaborate on the specific wavelengths contributing to the fit of this model. Oxidation increased absorption throughout the spectrum with particular effect between 4545-5000 cm^{-1} (2000-2200 nm), which was attributed to increased unsaturation and peroxide formation.

Cayuela Sanchez et al. (2013) also investigated conjugated dienes using K232 in virgin olive oil samples. Although transmittance was used in 5 mm path length quartz cuvettes, the best PLS model obtained had an RPD of 2.56, the worst value obtained among studies with such sample presentation apparatus in Table 6. This likely indicates the sensitivity of NIR for conjugated dienes as the study was a survey of collected oils, rather than an ASL, resulting in a range of K232 values approximately one quarter of the size used by Manley and Eberle (2006). Although the visible/NIR range of 350-2500 nm (4000-28,571cm^{-1}) was used initially, the model was the result of multiple rounds of wavelength selection based on regression coefficients. The final wavelengths used in this model were not reported.

Table 6. NIR models for conjugated dienes in lipids from the literature.

Reference	Lipid	Number of Samples	ASL Conditions	Sample Presentation	Range (% CD or K232)	Model	SEP	Calibration r^2	Validation r^2	RPD
Yildiz et al., 2001	Soybean Oil	128 (73 Calibration and 43 Validation from 16 Treatments of 8 Oils)	Fluorescent Light (12hr Intervals up to 180 hrs)	Transmittance in 1 mm path length Quartz Cuvettes	Range Not Specified Calibration (Mean 0.31% SD 0.09%) Validation (Mean 0.31% SD 0.11%)	PLS (9 Factors)	0.025% (0.021% for External Validation)	0.857	0.839 (0.912 for External Validation)	4.4 (5.2 for External Validation)
						MLR (4λ) 2064 nm (4845 cm^{-1}) 2430 nm (4115 cm^{-1}) 2350 nm (4255 cm^{-1}) 1398 nm (7153 cm^{-1})	0.023% (0.017% for External Validation)	0.837	0.856 (0.901 for External Validation)	4.8 (6.5 for External Validation)
				Transmittance in 2 mm path length Quartz Cuvettes		PLS (12 Factors)	0.020% (0.021% for External Validation)	0.927	0.893 (0.857 for External Validation)	5.5 (5.2 for External Validation)
						MLR (4λ) 2070 nm (4831 cm^{-1}) 1164 nm (8591 cm^{-1}) 1396 nm (7163 cm^{-1}) 1406 nm (7112 cm^{-1})	0.022% (0.018% for External Validation)	0.824	0.865 (0.882 for External Validation)	5.0 (6.1 for External Validation)
Manley & Eberle, 2006	Extra Virgin Olive Oil	104 (70 Calibration and 34 Validation from Treatment of 23 Oils)	35°C With Illumination for up to 10 Weeks	Transmittance in 0.2 mm Quartz Cuvettes	K232 1.67 - 20.36 Calibration (Mean 4.76 SD 3.68) Validation (Mean 4.61 SD 3.40)	PLS (6 Factors)	0.94		0.94	3.62
				Transflectance using glass Petri dishes with 0.3 mm high reflector		PLS (4 Factors)	1.33		0.87	2.56
Cayuela-Sanchez et al., 2013	Virgin Olive Oil	278 (223 Calibration and 55 Validation)	NA (Not an ASL)	Transmittance in 5 mm Path Length Quartz Cuvettes	K232 0.9 - 5.0 Calibration (Mean 2.0 SD 0.8) Validation (Mean 2.0 SD 0.8)	PLS (Not Reported)	0.32	0.82		2.56

2.5.3 Carbonyls

Researchers have investigated the use of NIR to determine secondary products of lipid oxidation in both neat oils and food systems. These secondary products comprise a variety of carbonyl compounds, including saturated and unsaturated aldehydes, ketones and acids. Unfortunately, direct evaluation of carbonyls is not possible because these groups do not give rise to signals in the NIR region. (See, e.g., Yildiz et al., 2001; Dellarosa et al., 2015) Thus, NIR assays must rely on changes in related moieties such as aldehydic C-H bonds which do exhibit signals to determine secondary lipid oxidation products. Attempts to model carbonyl values, hexanal, and free fatty acids using NIR have met with mixed success.

2.5.3.1 Total Carbonyl Value

During their investigation of quality alterations in lard upon extended heating at frying temperatures ranging from 160-230 °C, Szabo et al. (2009) also assessed the ability of NIR to model carbonyl values as determined by hydroxylamine-HCl titration. Carbonyl values of the 33 samples ranged from 3.61-21.6 with a mean of 8.18 and standard deviation of 4.59. The best PLS model was based upon a single factor and arose from a very weak calibration ($r^2 = 0.109$). The standard error of cross validation (SECV) reported (4.57) indicated a very poor RPD of 1.00 for the model. The authors speculated that very low concentrations of carbonyls could be responsible for the lack of correlation with NIR.

2.5.3.2 Hexanal

NIR has been determined to be moderately predictive of hexanal content in walnut kernels (Jensen et al., 2001) as well as peanuts and muesli (Jensen et al., 2004) (Table 7). These studies relied upon static GC headspace analysis as the reference assay for hexanal, reported in mg per kg of food product. The nuts were investigated as high fat foods (walnuts 53% fat w/w; peanuts 50% fat w/w), while muesli products were high carbohydrate foods (62 and 65% w/w for products I and II). Peanuts were evaluated at five time points within their shelf life as well as one

at the twenty four week shelf life and one at twenty six weeks (Jensen et al., 2004). Both muesli products were evaluated at four, seventeen and twenty six weeks, well within their thirty six week shelf life. However, muesli I was exposed to 2900 lux fluorescent light, while muesli II was further investigated at thirty four weeks and considerably beyond its shelf life at thirty nine, forty two and fifty two weeks (Jensen et al., 2005).

PLS models of hexanal using the full visible/NIR spectral range had moderate correlation with validation samples in walnuts ($r^2 = 0.72$), peanuts ($r^2 = 0.64$), muesli I ($r^2 = 0.70$), and muesli II ($r^2 = 0.83$) (Table 7). While limitation of the spectral ranges used to certain areas of the NIR did not improve models in nuts, validation statistics for samples of muesli I improved ($r^2 = 0.80$) by use of 5650 - 5994 cm^{-1}. Thus, it appears hexanal models could be slightly better in high carbohydrate foods rather than high fat ones. Oddly, the best model in Table 7 was obtained for muesli II, which had the smallest range of hexanal values among samples. This may indicate an issue with these results. The literature indicates that moderate correlation of hexanal content can be made using NIR. Jensen et al. (2004) concluded that spectroscopic methods were useful to complement though not replace chemical assays.

2.5.3.3 Free Fatty Acids

Attempts to use NIR to model free fatty acids (FFA), usually as percentage or concentration of oleic acid content, have been moderately successful in both neat lipid systems and solid food systems. These models arose from investigations of shelf life studies of neat lipids (Cho et al., 1998; Szabo et al., 2009), discrimination of plant oils from various sources (Manley & Eberle, 2006; Armenta et al., 2007; Cayuela Sanchez et al., 2013), discrimination of oil byproduct of fish meal (Cozzolino et al., 2005) and fish meat (Karlsdottir et al., 2014) by species and collection season, and discrimination of beef (Realini et al., 2004) and poultry breast meat (Berzaghi et al., 2005) based on dietary conditions.

Table 7. NIR models of hexanal content from the literature.

Reference	Food System	Number of Samples	ASL Conditions	Sample Presentation	Range Hexanal (mg/kg)	Model (Factors)	Spectral Range	RMSECV (mg/kg)	Validation r^2
Jensen et al., 2001	Walnut Kernels	50 (2 Samples Each at T0 and 6 Time Points for Each of 4 Conditions)	Storage in Light or Darkness at 5°C or 21°C for up to 25 Weeks	Reflectance in Rotating Sample Cup with Quartz Window and Compressive Paper Disk	0 - 202.1 (Mean 29.9 SD 49.3)	PLS (5)	400 - 2498 nm (4003 - 25000 cm^{-1})	26.2	0.72
						PLS (3)	1850 - 1990 nm (5051 - 5405 cm^{-1})	27.8	0.69
Jensen et al., 2004	Peanuts	170 at 7 Time Points	27°C in Darkness with up to 21% Oxygen in Headspace and Low Oxygen Permeable Packaging for up to 26 Weeks	Reflectance in Foss NIRSystems Model 6500 using Small Ring Cup and Spinning Module (NR-6506)	0.0 - 7.3 (Mean 1.3)	PLS (11)	400 - 2500 nm (4000 - 25000 cm^{-1})	0.7	0.64
		168 at 7 Time Points				PLS (6)	930 - 1034 nm (9671 - 10,753 cm^{-1})	0.7	0.58
	Muesli Product I	58 at 3 Time Points	27°C with Exposure to Fluorescent Light with up to 21% Oxygen in Headspace and Low or High Oxygen Permeable Packaging for up to 26 Weeks		0.9 - 141.9 (Mean 20.5)	PLS (4)	400 - 2500 nm (4000 - 25000 cm^{-1})	1	0.70
		60 at 3 Time Points				PLS (4)	1668 - 1770 nm (5650 - 5995 cm^{-1})	0.8	0.80
	Muesli Product II	167 at 7 Time Points	27°C in Darkness with up to 21% Oxygen in Headspace and Low or High Oxygen Permeable Packaging for up to 52 Weeks		0.0 - 1.5 (Mean 0.7)	PLS (8)	400 - 2500 nm (4000 - 25000 cm^{-1})	0.2	0.83

2.6 **Issues raised by analysis of the literature**

Considering all the studies cited above, issues of concern that impede the progress and acceptance of NIR as a valid and useful tool for analyzing lipid oxidation include the following:

1) Most calibrators are prepared from lipids extracted from foods and analyzed for lipid oxidation while NIR analyzes intact food samples with matrices intact and lipids unmodified.

2) Non-homogeneity of foods on a molecular scale leads to large variation in NIR data, which in turn impairs ability to develop and apply accurate mathematical models of foods.

3) Chemical analyses used to quantitate target molecules in calibrators seldom have the same sensitivity as NIR, impairing strong correlations in mathematical models.

4) Detailed analysis of chemometric models is often overlooked, leading to the possibility of incorrect correlation of spectral features with chemical analyses.

Much still needs to be learned to optimize NIR analyses of solid samples, particularly in the context of lipid oxidation in foods.

3. HYPOTHESIS AND OBJECTIVES

3.1 Hypothesis and Overall Objective

Fourier Transform Near Infrared Spectroscopy and Chemometrics can provide qualitative and quantitative information about lipid oxidation in complex food systems with minimal sample preparation, as reflected by using a solid food model system having minor amounts of oils analyzed for conjugated dienes, lipid hydroperoxides, and carbonyl products of lipid oxidation.

Many food companies have rejected use of NIR for on-line and quality control analyses of lipid oxidation, claiming poor sensitivity, large data scatter, and general inaccuracy, i.e. factors inherent in NIR as a methodology. Some basic research has also reported poor performance of NIR in identifying lipid oxidation in foods or model systems. This research questions these assessments and seeks to determine if apparent problems with NIR analyses stem from limitations of NIR spectroscopy itself or rather from practical issues of sample handling and analysis.

Accordingly, we start with the working hypothesis that NIR indeed has the capabilities to detect and differentiate products of lipid oxidation in foods and model systems, and propose that poor results with NIR analyses arise from suboptimal handling of samples, inappropriate or inaccurate calibrators, and random application of chemometric statistical analyses. Development of robust and accurate NIR analyses will probably require considerable revision and tailoring of approaches, rather than just transferring methodologies, e.g. from mid-IR.

3.2 Specific Objectives

1. Identify and minimize sources of spectral variation to improve methods for sample handling and presentation to the FT-NIR and enable analysis of lipid oxidation in complex solid food model systems using sample sizes practical for a research laboratory.

a. Test effects of compacting solid food samples rather than analyzing loose particles, and of forming pressure used to compact miniaturized samples of solid food model systems of oil mixed with white rice flour on variability of FT-NIR spectra and principal component scores.

b. Test effects of rotational averaging in miniaturized samples of solid food model systems of oil mixed with white rice flour on variability of FT-NIR spectra and principal component scores.

2. Test the ability of FT-NIR to discriminate among fats and oils present as a minor ingredient in a model solid food system primarily composed of white rice flour.

3. Test the ability of FT-NIR to discriminate among model solid food systems with different oxidation levels in oils present as a minor ingredient by preparing the model systems with fresh soybean, canola, sunflower or safflower oil as well as with oils subjected to oxidation at 60°C for one, two or three weeks.

4. Test the ability of FT-NIR to accurately quantify three markers of lipid oxidation, conjugated dienes, lipid hydroperoxides, and carbonyls, in solid food model systems with different oxidation levels in lipids present as a minor ingredient, by including canola and pecan oils chemically assayed for those markers both fresh and after oxidation at 40°C at six time points up to fifteen weeks.

4. EXPERIMENTAL PROCEDURES

4.1 Experimental Design

This study was undertaken to evaluate the performance of FT-NIR for analysis of lipid oxidation in solid food systems and to determine factors impacting that performance. Our laboratory recently acquired an FT-NIR instrument with a sample platform dedicated for the analysis of solids, thus we needed to prepare a model solid food system. The system was chosen to model the composition of cereal-based food products. White rice flour was chosen as the main constituent based on its widespread use in food products. Given the complexity imparted to the system by food processing techniques and the need for certainty that lipid oxidation per se was being modeled, shelf life studies were conducted by exposing neat lipids to oxidation at elevated temperatures prior to mixing with white rice flour at room temperature at the time of analysis. This procedure eliminated the need to extract lipids from a solid matrix prior to chemical analysis and thus provided lipids of identical composition for both chemical and NIR analyses. Preliminary investigations using moderate sample sizes indicated a scattering problem in spectra. Thus, the overall experimental design addressed the development and testing of a miniaturized FT-NIR sample platform prior to conducting an accelerated shelf life (ASL) study to assess the ability of FT-NIR to quantitatively monitor lipid oxidation in model solid food systems.

4.1.1 Development of a Miniaturized FT-NIR Sample Platform

FT-NIR models of chemical or physical parameters require a large number of replicates to ensure accuracy and fit. Depending upon the sample presentation format available, this can give rise to the need for large amounts of sample as well as related requirements for glassware and incubator and storage space. This makes the use of FT-NIR for systematic lab-scale investigation prohibitive without an appropriate mini or microscale sample presentation format. Thus, we developed a miniaturized format for use in the shelf life study reported herein as well as for general analyses of foods and model systems by NIR.

FT-NIR analyses of complex solid food systems are complicated by variation due to inhomogeneity of composition and differences in sample packing. Variation can be particularly confounding in analyses involving only minor changes in composition among different samples. The shelf life study herein was one such case, as changes were limited to those attributable to oxidative degradation of a minor constituent of the model system. Accordingly, we assessed means to reduce spectral variation by controlling material packing and increasing spectral sampling by rotation in a miniaturized sample holder. Successful modifications should improve modeling not only of the markers of lipid oxidation, but also of minor constituents in food systems more generally.

The experimental design for development and testing of the FT-NIR miniaturized sampling system is shown in Figure 10. Studies were undertaken to determine if FT-NIR data acquired from small samples was reproducible and could provide models which accurately discriminate among food systems containing different lipids. Additionally, studies were undertaken to determine if FT-NIR could discriminate among small samples made with the same lipid at different stages of oxidation. Finally, studies were undertaken to determine the effects of forming pressure applied during sample preparation and of sample rotation during FT-NIR analysis on variation in scores plots. These variation studies used models based on pressure differences per se as well as the best model obtained for qualitative discrimination of oxidation in canola oil mixed with white rice flour. The latter provided insight into the degree to which pressure and rotation could affect the results of FT-NIR analyses of lipid oxidation.

4.1.2 Assessment of FT-NIR for Quantitation of Indicators of Lipid Oxidation

The experimental design for the accelerated shelf life study to determine the ability of FT-NIR to quantitatively monitor lipid oxidation in canola and pecan oils incorporated in model solid food systems is shown in Figure 11.

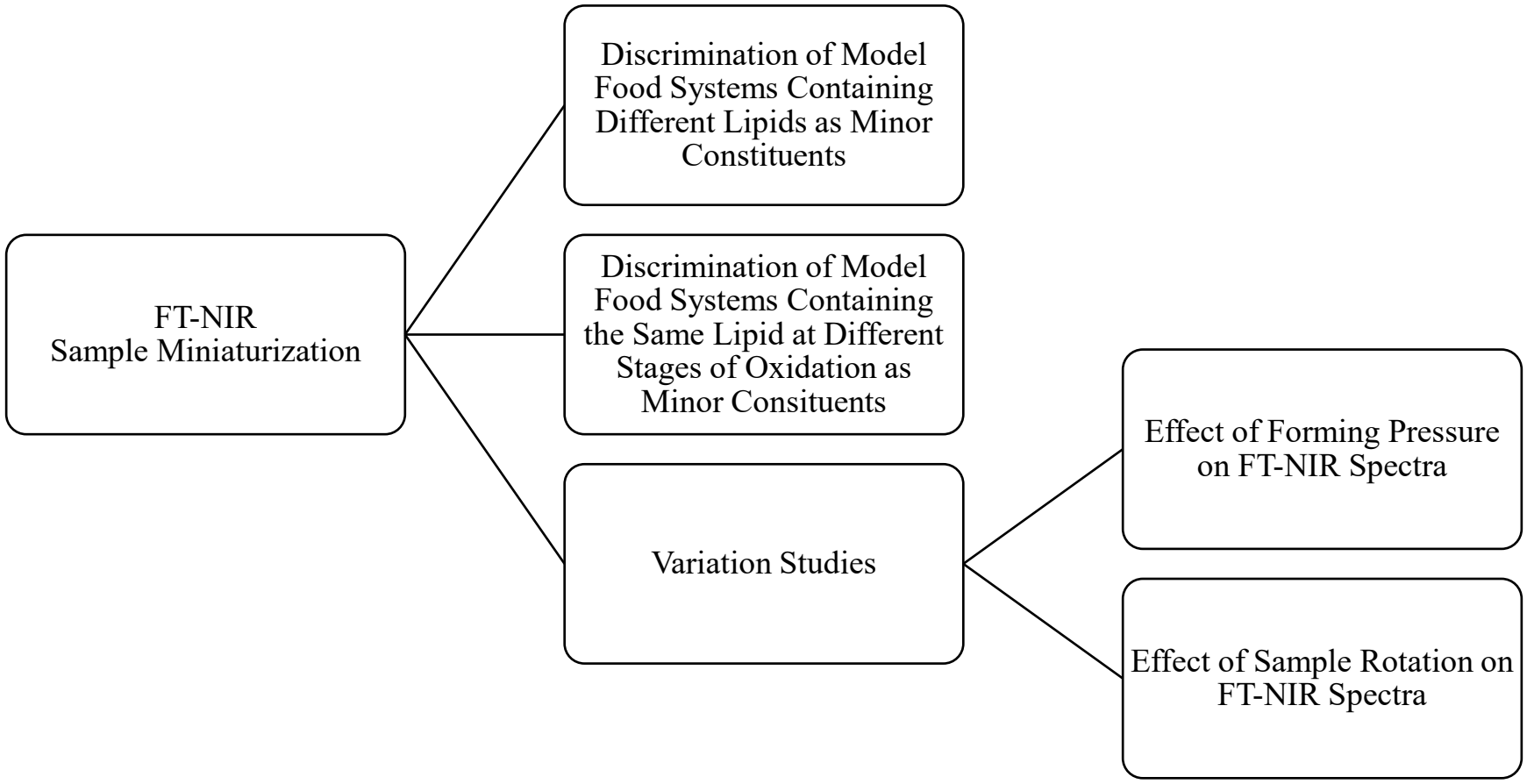


Figure 10. Experimental design for preliminary investigation of FT-NIR analysis of model solid food systems.

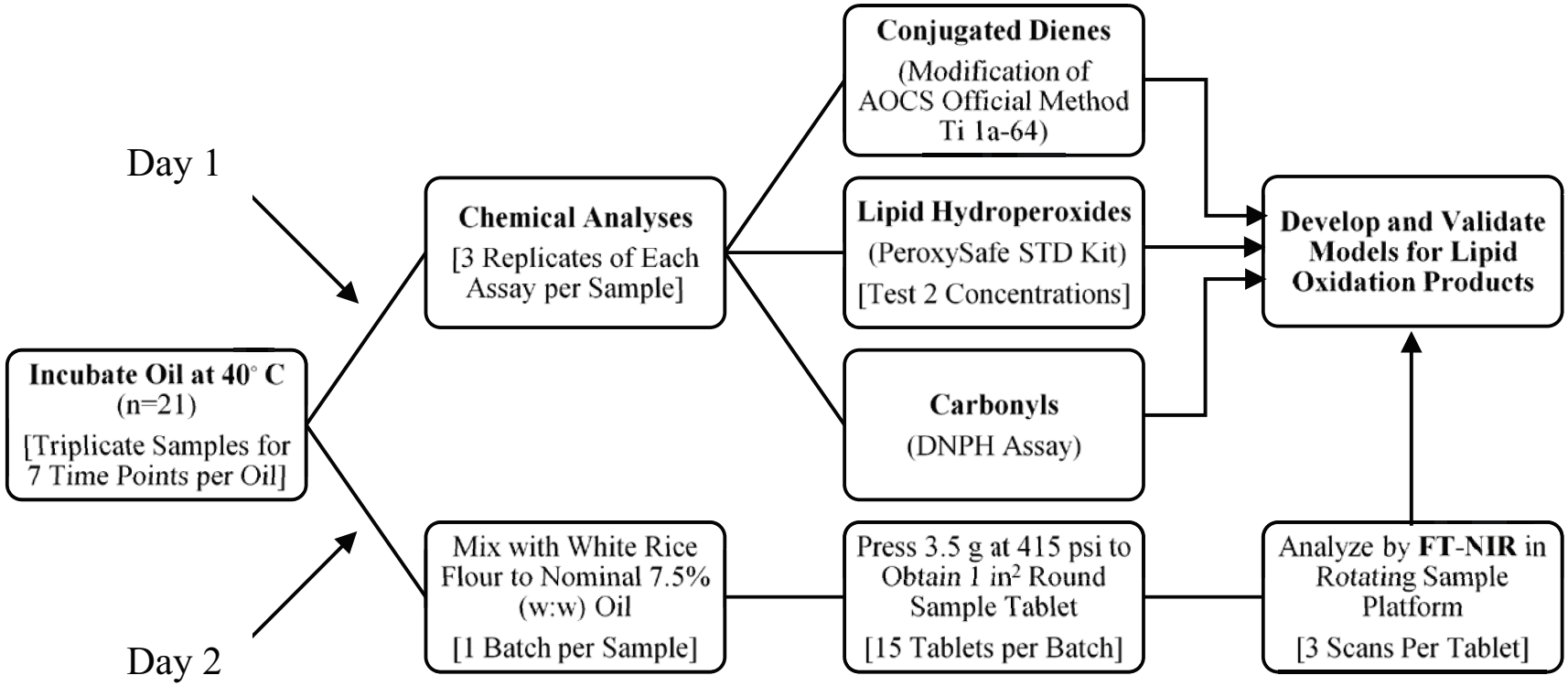


Figure 11. Experimental design for accelerated shelf life studies of pecan and canola oils.

Neat oils were incubated at 40°C to induce oxidation and triplicate samples were withdrawn at different time points for chemical analysis of three markers of lipid oxidation, namely, conjugated dienes, lipid hydroperoxides and soluble carbonyls. The same oil samples were mixed with white rice flour for FT-NIR analysis on the following day using the miniaturized sampling system developed herein. Data from reference and spectroscopic assays were then used to develop and test models for each marker in the respective model solid food systems.

4.2 Materials

4.2.1 Solvents

18 MΩ resistivity water was obtained by purification of doubly-distilled deionized water through a four-cartridge Milli-Q™ water purification system with two ion exchange cartridges for removal of metals. (EMD Millipore Corporation, Billerica, MA) This purified water is referred to as distilled deionized water herein.

Acetonitrile (Omnisolv® LC-MS Grade) was acquired from EMD Millipore (Billerica, MA, USA).

Chloroform (J.T. Baker HPLC Grade), Isooctane (2, 2, 4-Trimethylpentane) (Macron Fine Chemicals™ ChromAR®) and Isopropyl Alcohol (Macron Fine Chemicals™ UltimAR® HPLC Grade) were acquired from Avantor Performance Materials (Center Valley, PA, USA).

N, N – Dimethylformamide (Chromasolv® Plus for HPLC (≥99.9%)) was acquired from Sigma-Aldrich (St. Louis, MO, USA).

4.2.2 Reagents

2, 4 – Dinitrophenylhydrazine (70%) was acquired from Aldrich Chemical Co., Inc. (Milwaukee, WI, USA).

Formic Acid (≥98%) was acquired from EMD Millipore (Billerica, MA, USA).

Lauric Aldehyde (≥98%) was acquired from Aldrich (St. Louis, MO, USA).

The PeroxySafe™ STD Kit was acquired from MP Biomedicals, LLC (Solon, OH, USA).

4.2.3 Sample Materials

White Rice Flour (Remyflo R 7-150) was acquired from Beneo Inc. (Morris Plains, NJ, USA).

Pecan Oil was received as a generous donation from Kinloch Plantation Products (Winnsboro, LA, USA).

Chicken Fat was received as a generous donation from Royal Canin U.S.A., Inc. (Saint Charles, MO, USA).

The following fats and oils were purchased in local stores:

Canola Oil, Wegmans, Wegmans Food Markets (Rochester, NY, USA).

Flaxseed Oil, GNC Certified Organic, General Nutrition Corporation (Pittsburgh, PA, USA).

Lard, Leidy's®, ALL Holding Co., Inc. (Harleysville, PA, USA).

Palm Oil, Spectrum® Naturals Organic Shortening, Hain Celestial Group, Inc. (Lake Success, NY, USA).

Safflower Oil, Hollywood®, Hain Celestial Group, Inc. (Boulder, CO, USA).

Soybean Oil, Crisco®, The J.M. Smucker Company (Orrville, OH, USA).

Sunflower Oil, Loriva®, Blue Marble Brands, LLC (Providence, RI, USA).

4.3 Equipment

The Buchi NIRFlex N500 FT-NIR (Buchi Corporation, New Castle, DE, USA) is a modular instrument for analyzing many types of materials. Our experiments were conducted with the transflectance-based solids module (Figure 12). Transmission involves passing radiation through a sample and detecting the modified beam on the opposite side. Reflectance involves detecting the modified beam bouncing back from the sample surface on the same side as the incident radiation. Transflectance, like reflectance, involves detecting the modified beam bouncing back from the sample; however, like transmittance, the radiation penetrates beyond the

sample surface. Accordingly, transflectance involves collecting radiation which interacts with the sample prior to returning to a detector placed on the same side of the sample as the source of the impinging radiation.

Figure 12. Buchi NIRFlex Solids N500 FT-NIR instrument used in this study.

The Buchi FT-NIR accommodates a number of platforms for sample presentation. The instrument was originally acquired with a rotating platform for large petri dishes full of solid materials. Unfortunately, the amount of sample as well as incubator and freezer space plus glassware required was excessive for a systematic research program, making analysis of a large number of samples difficult if not prohibitive. Accordingly, an adjustable vial holder was acquired to accommodate small samples. Initial experiments using borosilicate glass shell vials to hold small amounts of sample indicated a high degree of spectral variation due to inconsistent

thickness and poor optical quality of vial bottoms. More importantly, the holder did not rotate and thus allowed only single point measurements for each sample, which per the manufacturer were unadvisable for the powdered sample mixtures of our model food systems.

Also, a bed of powdered or granular sample contains gaps of air at random locations, which can result in the detection of NIR radiation that has not fully interacted with the sample. Such non-interacted radiation has been reported to lead to spectral variation (Yoon et al., 2013). Curvilinearity of response in NIR measurement of powdered samples is also known to give rise to variation (Barnes et al., 1989). Although certain pretreatments may be applied to raw spectra to reduce these effects, the sensitivity of our application for lipid oxidation analyses in complex food systems required better reproducibility than these algorithms could provide. We therefore sought to apply a forming pressure to the samples to minimize air dispersed within the samples.

Sample packing is known to affect NIR spectra to such an extent that NIR has been applied to measure tablet compaction in the pharmaceutical industry (Guo et al., 1999; Roggo et al., 2005). To test the effect of applying a forming pressure in our model systems, samples were pressed into tablets using a manually-operated Carver hydraulic press (Carver, Inc., Wabash, IN, USA) (Figure 13). Following initial experiments, the incorporation of a digital pressure gauge (Figure 13, bottom right) confirmed that the Carver was incapable of applying pressure reproducibly to each sample. Thus, it was necessary to ascertain whether the differences in pressure exerted on different samples during forming added to variability among sample spectra, confounding the NIR analysis.

To explore this issue, a lever press (Figure 14) was constructed to ensure application of reproducible pressures during forming of the samples. The lever press arm weighed 65 lbs and had holders positioned to confer a mechanical advantage of 5X, 10 X, 15 X and 20 X to any weights applied. Plate weights were obtained from a local fitness store and masses verified with a Toledo scale (Mettler-Toledo, Columbus, OH, USA).

Figure 13. Carver hydraulic press used to prepare sample tablets for NIR analysis.

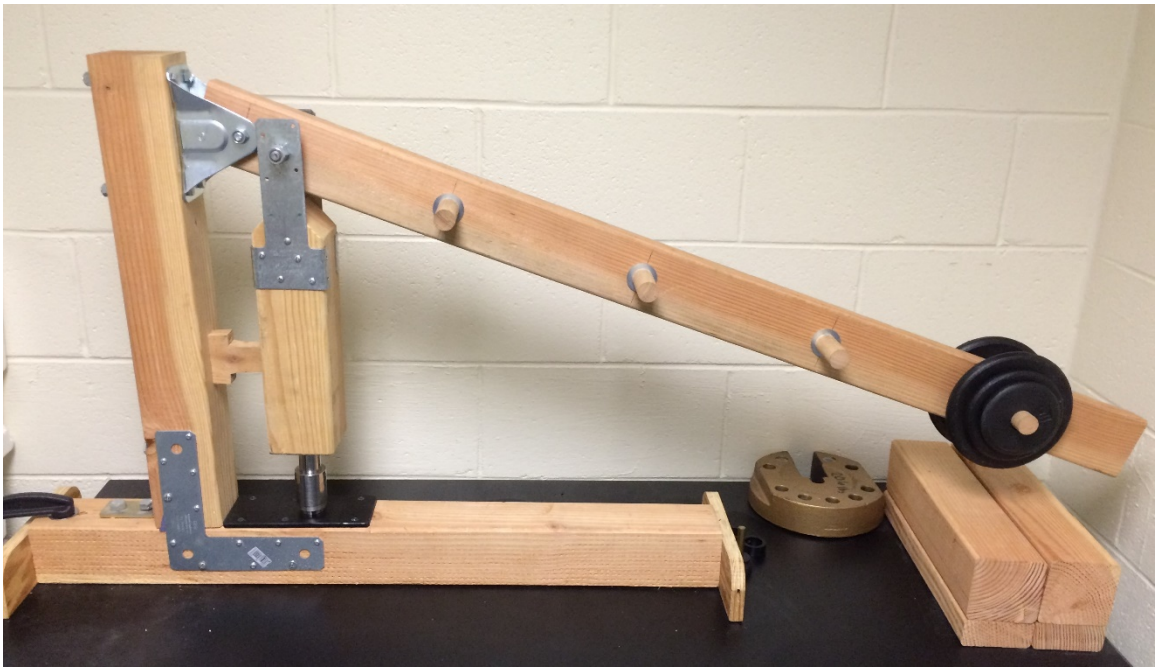


Figure 14. Lever press constructed for compacting samples at discrete pressures.

Approximately 4 g of each sample were pressed in a 1.125 inch diameter die which provided a 0.994 square inch sample tablet (Figure 15). In initial experiments, spectra were acquired by placing the pressed sample on a washer which was then inserted directly over the coverplate window. This window was centrally disposed over the rectangular NIR beam, resulting in a stationary scan on the center of the tablet through the hole in the middle of the washer. To reduce variation by expanding the surface area scanned by the NIR, a rotating sample holder was mounted on the original coverplate as shown in Figure 16. The motorized apparatus turned a sample holder insert one revolution every 15 seconds (4 rpm).

Figure 15. Compacted 7.5% (w/w) canola oil : white rice flour samples.

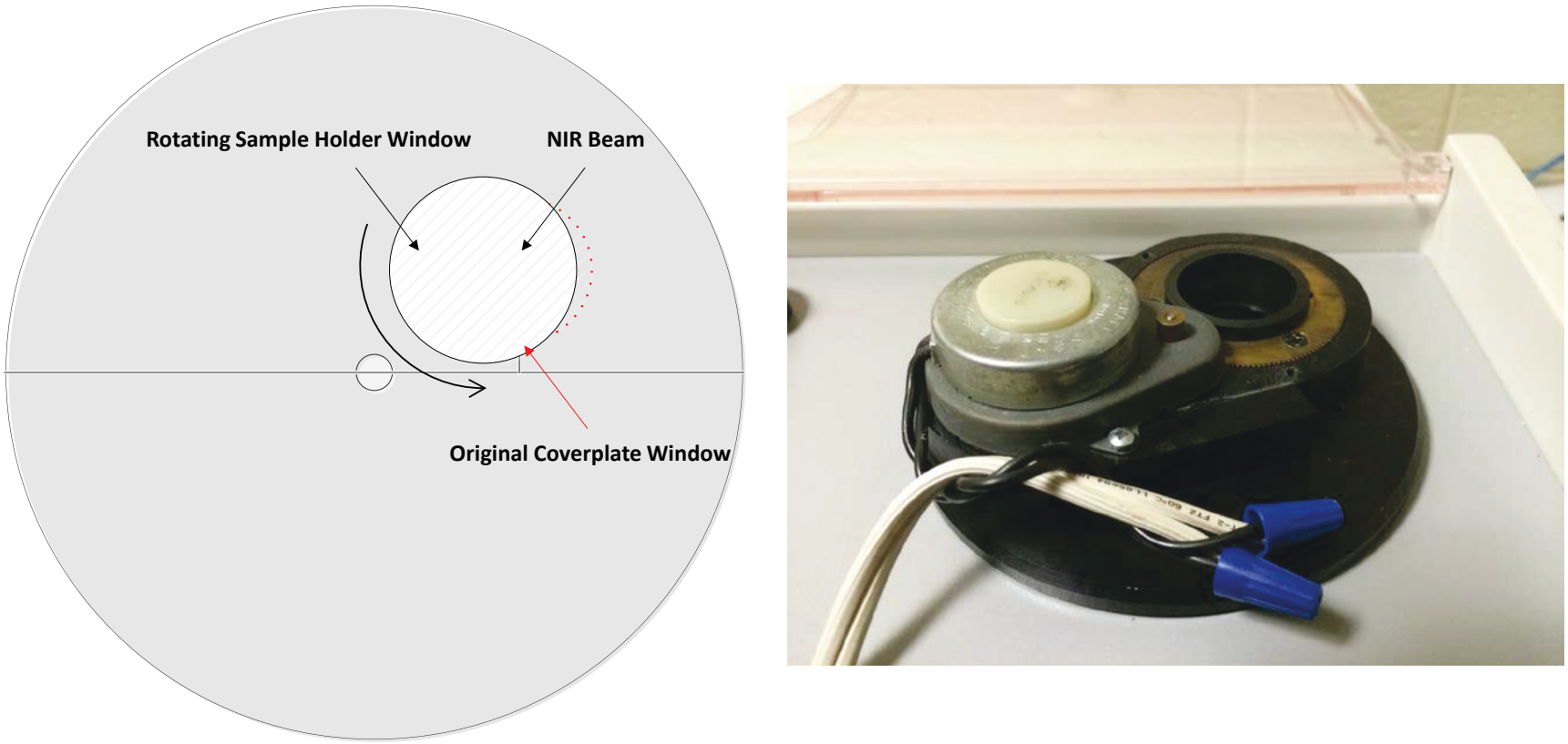


Figure 16. Schematic drawing (left) and photo (right) of rotating sample holder.

The sample holder was disposed off-center from the original coverplate window so that the beam would sweep over the full surface of the sample during rotation. Compacted samples varied in oil content from 7.5 to 15% (w:w), and ranged from moderately to very friable depending upon oil content and forming pressure applied. To structurally support the samples during measurement and accommodate scanning of the larger surface area, they were placed on 1.125 inch diameter, 2 mm thick borosilicate glass discs (Specialty Glass Products, Inc., Willow Grove, PA, USA) (Figure 17). The discs rested on a lip in the base of the sample holder insert. These pressing and sample rotation techniques were used in all quantitative FT-NIR analyses in this study.

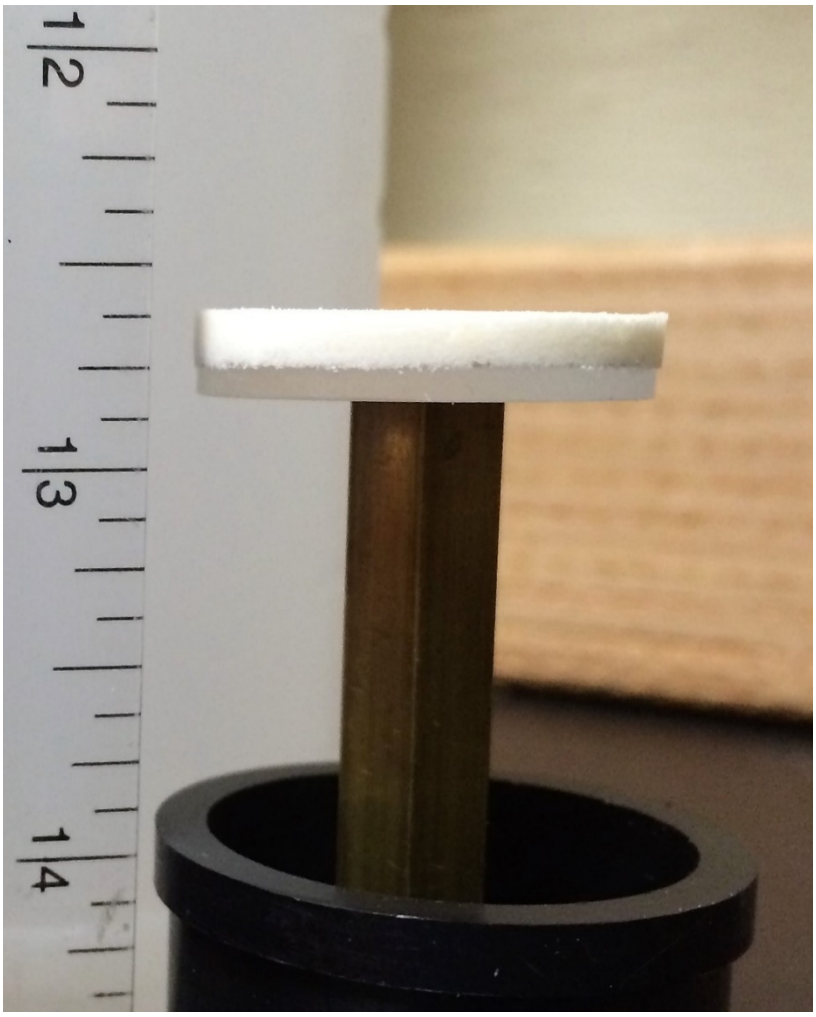


Figure 17. Insertion of compacted sample on borosilicate glass disc in sample holder.

4.4 Experimental Methods

4.4.1 Discrimination of 15% w/w Lipid : White Rice Flour Mixtures

Commercial products of canola oil, sunflower oil, soybean oil, safflower oil, flaxseed oil, chicken fat, palm oil, lard and white rice flour were obtained from the sources cited in 4.2 above. In stainless steel bowls, 61 g of each oil was coated onto 345 g of white rice flour using a handheld electric mixer to make mixtures of 15% (w/w) lipid: white rice flour. Ten samples were prepared from each mixture by pressing 4 g in a 1.125 inch diameter die using a Carver press at one half metric ton of pressure for three seconds. After 20 minutes, each sample was placed on a washer and three replicate spectra were scanned on the NIRFlex Solids N500 without rotation (Figure 18). Qualitative models to discriminate among the different lipid mixtures were generated in NIRCal by dedicating seven of the samples to calibration and three to validation sets, respectively.

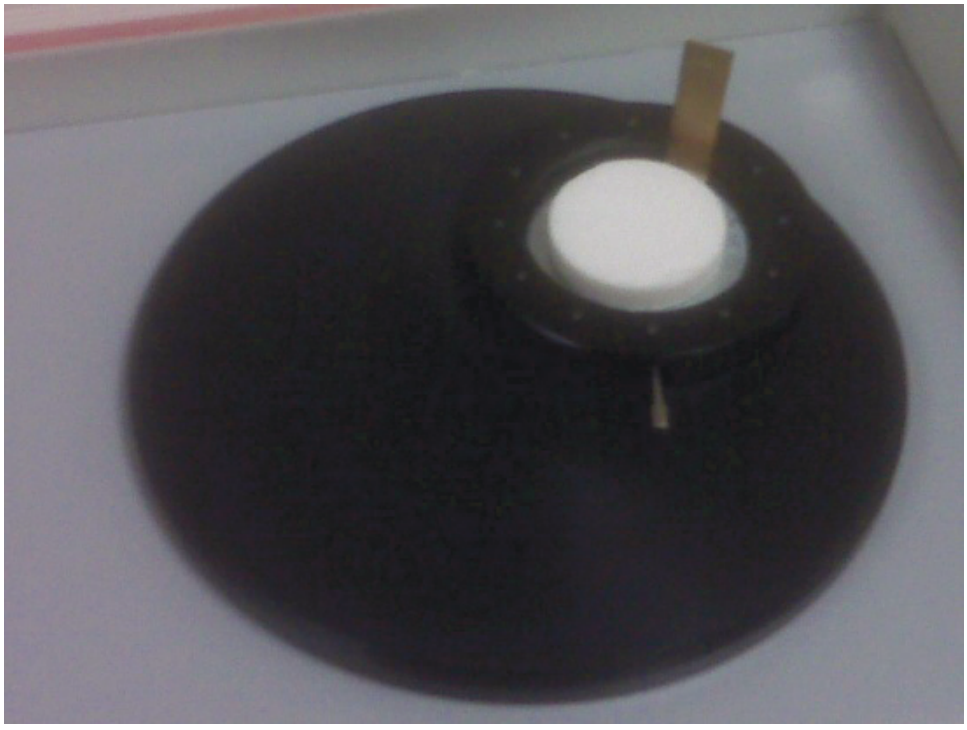


Figure 18. Scanning of a stationary sample tablet on a washer.

4.4.2 Discrimination of Degree of Lipid Oxidation : Qualitative ASL Study

To assess the effect of oxidative degradation of oils on the spectra of their respective mixtures with white rice flour, soybean oil, sunflower oil, safflower oil, and canola oil heated for up to three weeks at 60°C were assayed by NIR each week. Samples of 12 g of each oil were placed in 50 mL borosilicate glass serum bottles (Wheaton Industries, Inc., Millville, NJ, USA) and incubated in a Hybaid HS9360 rotisserie hybridization oven (Thermo Fisher Scientific, Waltham, MA, USA) at 60°C in darkness using a rotation speed setting of 6.5. A laboratory tissue was secured over the mouth of each bottle to block particulate matter while allowing oxygen to interact with the sample. Each week a single bottle of each oil was assayed, providing 4 time points including the T0 sample.

Each sample was withdrawn from the oven and allowed to cool to room temperature before 10 g of oil was coated on 56.67 g of white rice flour to produce a 15% (w:w) lipid mixture in a stainless steel mixing bowl using a whisk. Ten replicate samples were prepared for NIR by pressing 4 g of the mixture in a 1.125 inch diameter die using a Carver press at one half metric ton of pressure for three seconds. After one hour, each sample was placed on a washer and three replicate spectra were scanned on the NIRFlex NIR Solids without rotation (Figure 18). Qualitative models to determine if spectral features corresponding to oxidative changes in oils were discernible were generated in NIRCal by dedicating seven of the samples to calibration and three to validation sets, respectively.

4.4.3 Discrimination of Identical Samples Subjected to Different Pressures

In initial experiments, the Carver press was used to apply a “forming” pressure to remove air in an attempt to reduce spectral variability from packing differences inherent in powdered samples. However, addition of a precise digital pressure gauge to the press indicated this press could not be used to apply pressure reproducibly to each sample. Thus, it was necessary to ascertain whether the differences in pressure exerted on different samples during forming added

to variability among sample spectra, confounding the NIR analysis. The objective of this experiment was to determine whether NIR could discriminate among samples packed at various forming pressures far lower than those employed in pharmaceutical tableting operations where the use of NIR to monitor compaction has been reported (Guo et al., 1999; Roggo et al., 2005).

A 15% (w/w) mixture of canola oil and white rice flour was prepared by dropwise addition of 40 g of canola oil during mixing of 226.7 g of white rice flour in a kitchen stand mixer fitted with a dough hook attachment. Four sets of ten samples each were prepared by pressing 4 g samples in a 1.125 inch diameter die using the lever press with 13.75 lbs, 15 lbs, 16.25 lbs and 17.5 lbs applied to the 20X holder. Factoring in the die area of 0.994 in^2 and the weight of the lever arm, the nominal pressures applied to the respective sample sets were 340 psi, 365 psi, 390 psi and 415 psi. After one hour, each sample was placed on a washer and three replicate spectra were scanned on the NIRFlex NIR Solids without rotation (Figure 18). Qualitative models to discriminate among samples prepared with different forming pressures were generated in NIRCal by dedicating seven of the samples to calibration and three to validation sets, respectively.

This experiment was subsequently repeated using half the amount of oil when results of the original experiment indicated the highest pressure applied caused oil to separate from the flour and pool in the tablets, giving rise to increased inhomogeneity. Samples were made from a 7.5% (w/w) mixture of canola oil and white rice flour prepared by dropwise addition of 30 g of canola oil during mixing of 370 g of white rice flour in a kitchen stand mixer fitted with a dough hook attachment. Qualitative models to discriminate among samples prepared with different forming pressures were generated in NIRCal by dedicating seven of the samples to calibration and three to validation sets, respectively.

4.4.4 Effect of Sample Rotation

Two experiments were performed to determine the effect of rotation on the dispersion of scores derived from spectra using the qualitative canola ASL model generated from the

experiment described in 4.4.2.2 above. The miniaturized sampling system with the modified rotating coverplate described herein was used in both of these experiments.

A 10% (w/w) mixture of canola oil and white rice flour was prepared by dropwise addition of 72.2 g of canola oil during mixing of 650 g of white rice flour in a kitchen stand mixer fitted with a dough hook attachment. Each sample tablet took longer to monitor via NIR in these experiments since spectra were acquired with the same samples stationary and rotating. Thus, each tablet was prepared immediately before it was assayed. The sample mixture was placed in a sealed glass jar protected from light during the rotation experiments.

4.4.4.1.1 Replicate Scans of a Single Sample

The objective of this experiment was to assess the effects of inhomogeneity in a sample on the variability of scores as well as the potential effects of forming pressure and rotation on that variability. Replicate scans were made on a 4 g sample of the 10% canola oil: white rice flour mixture as a free powder and subsequently on another 4 g sample of the mixture as a tablet compressed at 415 psi. Twenty one replicate scans were acquired at different locations in the sample by positioning it over the detector using landmarks on the rotating sample holder. These scans comprised seven cycles of triplicates acquired in chronological order. In each cycle, each of the triplicate scans was acquired at a location offset by 120° angles relative to the other two. At each point, a stationary scan was acquired on the sample immediately prior to a rotating scan beginning at that point. Each spectrum took thirty two seconds to acquire, representing the average of sixty four scans acquired each half second, regardless of whether the sample was stationary or rotating during acquisition. Including the time for each scan as well as the time required to position the sample, the overall time of exposure for the sample on the NIR was just under half an hour. Sample spectra were translated to scores along the two principal component axes from the qualitative canola ASL model for comparative statistical analyses of the dispersion among replicates of stationary and rotating samples of powders and pressed tablets.

4.4.4.1.2 Replicate Scans of Multiple Samples

The objective of this experiment was to assess the effects of inhomogeneity among multiple nominally identical samples on the variability of scores as well as the potential effects of forming pressure and rotation on that variability. Replicate scans were made on two sets of ten samples of the 10% canola oil: white rice flour mixture. Each sample was made by pressing 4 g of the mixture into a 1.125 inch diameter tablet. The first set was gently compressed at 115 psi, while the second set was compressed at 415 psi. For each sample tablet three replicate scans were acquired, each at a location offset by 120° angles relative to the other two, by positioning the tablet over the detector using landmarks on the rotating sample holder. At each point, a stationary scan was acquired immediately prior to a rotating scan beginning at that point. Each spectrum took thirty two seconds to acquire, representing the average of sixty four scans acquired each half second, regardless of whether the sample was stationary or rotating during acquisition. Including the time for each scan as well as the time required to position the sample, the overall time of exposure for the sample tablet on the NIR was just under five minutes. Sample spectra were translated to scores along the two principal component axes from the qualitative canola ASL model for comparative statistical analyses of the dispersion among replicates of stationary and rotating sample tablets formed at low and high pressures.

4.4.5 Accelerated Shelf Life Study: Quantitative Analysis of Lipid Oxidation

To assess the ability of NIR to discern indicators of lipid oxidation, an accelerated shelf life (ASL) study was conducted using two commercial oils (pecan and canola) incubated for up to 16 weeks at 40° C in darkness in a Jeio Tech incubator equipped with a rotating platform set at 45 rpm (Jeio Tech, Inc., Woburn, MA, USA). Replicate 10-12 g samples of oil were incubated in 50 ml borosilicate glass serum bottles (Wheaton Industries, Inc., Millville, NJ, USA) covered with a laboratory wipe to block particles but allow for oxygen to interact with the sample. For each oil,

18 samples were incubated, allowing assay of three replicate samples for each of six time points in addition to three replicate samples at T0.

NIR data were correlated with three chemical analyses for indicators of lipid oxidation: conjugated dienes and peroxides as two early indicators of oxidation, as well as carbonyls, a later indicator of oxidation. Chemical analyses were performed on samples the day they were removed from the incubator. Due to time constraints, NIR assays were typically performed on the following day.

At each time point, three bottles containing the same type of oil were removed from the incubator, flushed with argon and sealed after withdrawal of an approximately 1 mL aliquot for use in chemical assays before stoppering the serum bottles, wrapping the closure with gas impermeable Teflon tape and storing under foil in the -80° C freezer. Each aliquot for chemical assays was kept under argon in a sealed borosilicate glass test tube in darkness. Throughout all chemical assays precautions were taken to protect the aliquots and assay samples from light and oxygen where possible.

4.4.5.1 Conjugated Dienes

Oils were assayed for conjugated dienes using a modified version of AOCS Official Method Ti 1a-64. Given the instability of these compounds, this assay was always performed first after samples were withdrawn from the incubator and equilibrated to room temperature given the instability of these compounds. 250 ml of isooctane was sparged with argon for at least 15 minutes and its absorbance at 233 nm was verified to be no greater than 0.070 against a distilled deionized water blank. Typically, two serial dilutions of each sample of oil in isooctane were made, with the second made directly in the microcuvette in which the absorbance reading was made. For time points after 10 weeks of incubation, it was necessary to incorporate a third serial dilution to account for the increased content of conjugated dienes. For each sample, three replicate dilutions were performed and the average result was used for regression of the NIR data.

Absorbance measurements were made with a Cary 50 Bio UV-Visible spectrophotometer (Varian, Inc., Palo Alto, CA, USA) using the Simple Reads program. Each replicate dilution was scanned three times for 3 seconds each at both 233 nm and 215 nm. The specific absorbance value or oxidation index (Klein, 1970) was obtained by dividing the average value for the sample at 233 nm by that at 215 nm, and normalized among all samples by multiplying by the dilution factor. The specific absorbance values were found to be much more stable among repeated measurements than absorbance values at 233 nm alone.

4.4.5.2 Lipid Hydroperoxides

Oils were analyzed for peroxide value using the PeroxySafe™ STD kit (MP Biomedicals, Solon, OH). Measurements were performed on a Cary 50 spectrometer using the Simple Reads program. The instrument was fitted with a platform to accept 10 mm borosilicate glass test tubes, which helped minimize use of the reagents while affording stable results by eliminating time-consuming transfers to cuvettes. Isopropyl alcohol was a major solvent in this reaction. Given the volatility of this solvent, opening the reaction vessel to transfer its contents to a cuvette gave rise to fluctuations in concentration which confounded results. Also, it was very difficult to use the same cuvette to read multiple samples given the time constraints of this assay discussed below. Accordingly, measurements were obtained in the same sealed vessel in which the reaction occurred.

Although this colorimetric assay is based on the ferrous xylenol orange assay which develops a stable color within 20 minutes to a half hour, the color in this kit continues to develop beyond the 15 minute incubation recommended by the manufacturer. Accordingly, this assay was found to be extremely time sensitive and it was necessary to read the sample exactly 15 minutes after addition of the final reagent to the reaction.

If the concentration of peroxides in a sample is too high, it can bleach the color complex and result in underestimation of the peroxide value. For this reason, it is advised to test at least two different dilutions of each sample to ensure accuracy of the result. On any given day, three

samples of the same oil incubated for the same length of time were tested. To address the bleaching concern, as well as the time sensitive nature of each analysis, each peroxide assay was performed using two sets of six samples each. The first set involved three analytical replicates of the first oil sample prepared at two different dilutions. This afforded the appropriate results for the first oil sample, as well as the appropriate dilution to use for the second and third samples. In the second set, three analytical replicates each of the second and third oil samples were prepared at the appropriate dilution and assayed.

The PeroxySafe™ reaction was prepared by flushing each 10 mm diameter test tube with argon prior to adding 1000 µL of Reagent A, 25 µL of the oil sample diluted as necessary in the Prep Reagent, and 100 µL of Reagent B. Upon addition of 160 µL of Reagent C, each sample tube was sealed, vortexed for 30 seconds and incubated in a test tube block at room temperature for precisely 15 minutes. All reactions were prepared and incubated in darkness. At exactly 15 minutes, each sample was scanned in the Cary 50 at 570 nm for three seconds. Peroxide values in milliequivalents per kilogram of oil (meq/kg) were determined by comparison with a standard curve generated the same week using calibrators provided in the kit.

4.4.5.3 Carbonyls

Oils were analyzed for carbonyls using an HPLC assay based on derivatization with 2, 4-dinitrophenylhydrazine (DNPH). Formic acid was chosen as the acidulant as it was volatile and thus would be compatible with LC-MS in the future. A pH of 3.0 was selected for the reaction medium, which struck a balance between the pH needs of the reagent and carbonyl substrates and rendered the sample suitable for direct injection without dilution on the HPLC column, which had a low-end pH limit of 2.5.

4.4.5.3.1 Recrystallization of DNPH

DNPH was recrystallized twice from n-butanol and once from acetonitrile to remove hydrazone impurities from opportunistic derivatization of volatile carbonyls in ambient air. To

recrystallize from n-butanol, 26 g of 70% DNPH was added to 200 ml of n-butanol in a 500 ml round bottom flask. A magnetic stirrer was added before the flask was fitted with a reflux condenser and placed in a heating mantle. The mixture was refluxed with stirring for one hour and then decanted while hot into a second round bottom flask for recrystallization and drying in a rotary evaporator (Buchi Corporation, New Castle, DE, USA) fitted with a water bath heated to 70°C. The recrystallized DNPH was dissolved in n-butanol and the process was repeated.

The twice recrystallized DNPH was then recrystallized from acetonitrile using the same procedure outlined above except that a 50°C bath was used with the rotary evaporator during recrystallization to avoid bumping because of the reduced boiling point of this solvent relative to n-butanol. Following the final recrystallization, the DNPH was dried in darkness in a vacuum oven at 50°C overnight. The final product was verified to be free of impurities by HPLC and stored in crystalline form at room temperature in darkness in a foil-wrapped borosilicate glass screw cap vial flushed with argon and sealed with gas impermeable Teflon tape.

4.4.5.3.2 Preparation of DNPH Reagent

DNPH reagent (202 mM / 10X reaction concentration) was routinely prepared on the day of each analysis by weighing 60 mg of recrystallized DNPH (M.W. = 198.14 g/mol) into an argon-flushed 12 mm diameter test tube, dissolving in 1.454 mL of dimethylformamide (DMF), and acidifying with 46.2 µL of 2.16 M formic acid. The latter was prepared by adding 247 µl of concentrated formic acid (26.24 M determined using d = 1.22 g/ml, M.W. = 46.03 g/mol, purity = 99%) to 2.753 ml of dd H_2O. The DNPH reagent was prepared in dark conditions and likewise kept in a sealed glass tube under argon at room temperature pending use within a few hours.

4.4.5.3.3 DNPH Reaction and HPLC Assay

The DNPH reaction was performed by combining 725 µl of pre-mixed 1:1 (v/v) isopropanol : acetonitrile (IPA:ACN), 125 µl of chloroform, 25 µl of the lipid sample, 25 µl of 40 mM lauric aldehyde (as an internal standard) and 100 µl of fresh DNPH reagent in an HPLC vial

pre-flushed with argon. Each assay included three blank replicates which were prepared by combining 750 µL of 1:1 (v/v) acetonitrile/isopropanol, 125 µL of chloroform, 25 µL of 40 mM lauric aldehyde, and 100 µL of DNPH reagent. The overall DNPH concentration in the reaction, including protonated (inactive) and free reagent, was 20.2 mM. All twelve reactions (three blanks and three replicates each of the three oil samples assayed that day) were prepared and incubated at room temperature in darkness for five hours prior to initiation of the HPLC run on the first sample.

HPLC was performed according to the conditions of Yao (Yao, 2015). The column was a Restek® Ultra C18 (4.6 mm I.D. x 150 mm length with 5 µm particle size), for which the manufacturer's recommended pH range was 2.5 – 8.0. The method used a gradient elution with two mobile phases, 1:1:2 (v/v) Isopropanol: Acetonitrile: ddH_2O (B) and 1:1 (v/v) Isopropanol: Acetonitrile (A), as shown in Table 8. The overall run time was 60 minutes.

Table 8. HPLC Gradient Conditions for DNPH Carbonyl Assay.

Time (min)	% B	Flow Rate (ml / min)
0	**83.3**	**1.2**
17	**0**	**1.2**
19	**0**	**1.4**
50	**0**	**1.4**
55	**83.3**	**1.2**
60	**83.3**	**1.2**

Analyses were performed using an Agilent 1100 Series liquid chromatograph equipped with a diode array detector. HPChemStation software provided system control as well as data analysis. In addition to 360 nm and 206 nm, wavelengths evaluated included 233 nm for monounsaturated carbonyls with π systems conjugated to the carbonyl bond (2-enals) and 270 nm

for polyunsaturated carbonyls such as decadienal with π systems conjugated to the carbonyl bond (2, 4-enals).

4.4.5.4 Near Infrared Spectroscopy

4.4.5.4.1 Data Collection

All NIR analyses were conducted on a Buchi NIRFlex N-500 NIR spectrometer with NIRWare 4.2 sample management and operating software and NIRCal 4.2 chemometric software (Buchi Corporation, New Castle, DE, USA). This instrument (Figure 12) works in transflectance mode, whereby the solid sample causes some of the incident radiation to be redirected back to the detector. For each sample in this quantitative oxidation study, a 7.5% w/w oil mixture was prepared by addition of 5 g of oil to 61.67 g of white rice flour (Remyflo R 7-150) (Beneo Inc., Morris Plains, NJ, USA) in a small stainless steel mixing bowl while mixing with a hand-operated OXO stainless steel egg beater (OXO International, Ltd., New York, NY, USA).

Mixing resulted in a heterogeneous product given the hydrophilic nature of rice flour. As the packing density of a powdered sample could influence the NIR spectra of samples, fifteen samples of 3.5 g each of the oil and flour mixture were pressed in a 1.125 inch diameter die under a uniform pressure of 415 psi using the lever press (Figure 14). Each sample tablet was covered with a 1.125 inch diameter borosilicate glass disc upon removal from the die. The tablet and glass disc were then inverted into a holder (Figure 17). The holder assembly mounted in a rotating sample platform (Figure 16) disposed off-center from the beam of near infrared radiation, allowing for the beam to sweep over the full area of the tablet as it rotated during spectral acquisition.

NIR spectra were acquired in triplicate on each of fifteen rotating samples. Each spectrum was a composite of 64 scans acquired once per half second (amounting to a 32 second scan time) at a resolution of 4 cm^{-1}. The forty five spectra acquired for each batch were averaged for use in chemometric analysis to create models for conjugated dienes and peroxide values using

the respective reference assay data. Chemometric approaches included principal component regression and partial least squares regression of spectra using a variety of pretreatments including normalization, smoothing and/or first or second derivatization of the original spectra. Cross validation was used to obtain the best model for each set of conditions from the software. Model quality was evaluated by the manufacturer's index, the Q value, described in Appendix B.

4.4.5.4.2 Assessment of Wavenumbers for Use in Models

It is possible to improve the quality of models by refining the range of wavenumbers employed therein. It is important to maintain as much of the valuable information contained in the spectrum as possible. However, inclusion of data from any given wavenumber adds noise in addition to any useful information it can provide. Thus, it is important to identify and exclude wavenumbers which lack sufficient useful information to overcome the additional noise they impart to a model.

4.4.5.4.2.1 Comparison of Spectral Variation Among and Within Groups: e

One metric used to assess the importance of wavenumbers for discriminating among samples is e, the ratio of variance between different groups of samples to the variance of samples within the same group. This is determined by the following formula (Wu et al., 1995):

$$e_i = \frac{\sum_{j=1}^{k} n_j \left(\bar{y}_{ji} - \bar{y}_{.i}\right)^2}{\sum_{j=1}^{k} \left(n_j - 1\right) s_{ji}^2}$$

The subscripts *i* and *j* denote the wavenumber and the sample group, respectively, while n_j indicates the number of samples in group *j*. In the numerator, $\bar{y}_{ji}$ is the mean absorbance of all samples from the same group at a given wavenumber, while $\bar{y}_{.i}$ is the overall mean absorbance at a given wavenumber for samples from all groups. In the denominator, s_{ji} is the standard deviation for all samples from the same group at a given wavenumber.

When a particular wavenumber is useful in discriminating among different types of samples, the sum of differences between the mean absorbance for each group and the overall

mean in the numerator will grow. When a particular wavenumber is stable among identical samples, the standard deviation for samples within the same group in the denominator will diminish. Wavenumbers which are both stable and useful for discrimination will thus have the largest values of e.

4.4.5.4.2.2 Cross Validation Regression Coefficients t-Test

The use of cross validation to estimate uncertainty in PLS regression was originally described by Martens and Martens (2000). This development enabled significance testing of PLS model parameters such as regression coefficients, scores and loadings, addressing a key criticism of PLS among statisticians at the time (Davies, 2001). Significance testing of regression coefficients was used to improve the quality of models by eliminating unimportant wavenumbers from models (Martens & Martens, 2000; Westad & Martens, 2000).

Cross validation is used to evaluate the performance of models in predicting unknowns (Williams, 2001; Naes et al., 2002). An example for models created with ten samples is shown in Figure 19. A test of the model is run for each sample by removing it from consideration during calibration for an analyte of interest. The prediction error for that test is determined by the difference between the amounts of analyte predicted by the NIR model ($\hat{y}$) and determined from the reference assay (y). Test results are combined in a single term, the Root Mean Square Error of Cross Validation (RMSECV), which decreases as predictive ability of the model improves. As the RMSECV is a composite of terms for the bias and standard deviation of residuals (see Appendix B) reported by the Buchi NIRCal 5.4 software, those terms were used here instead.

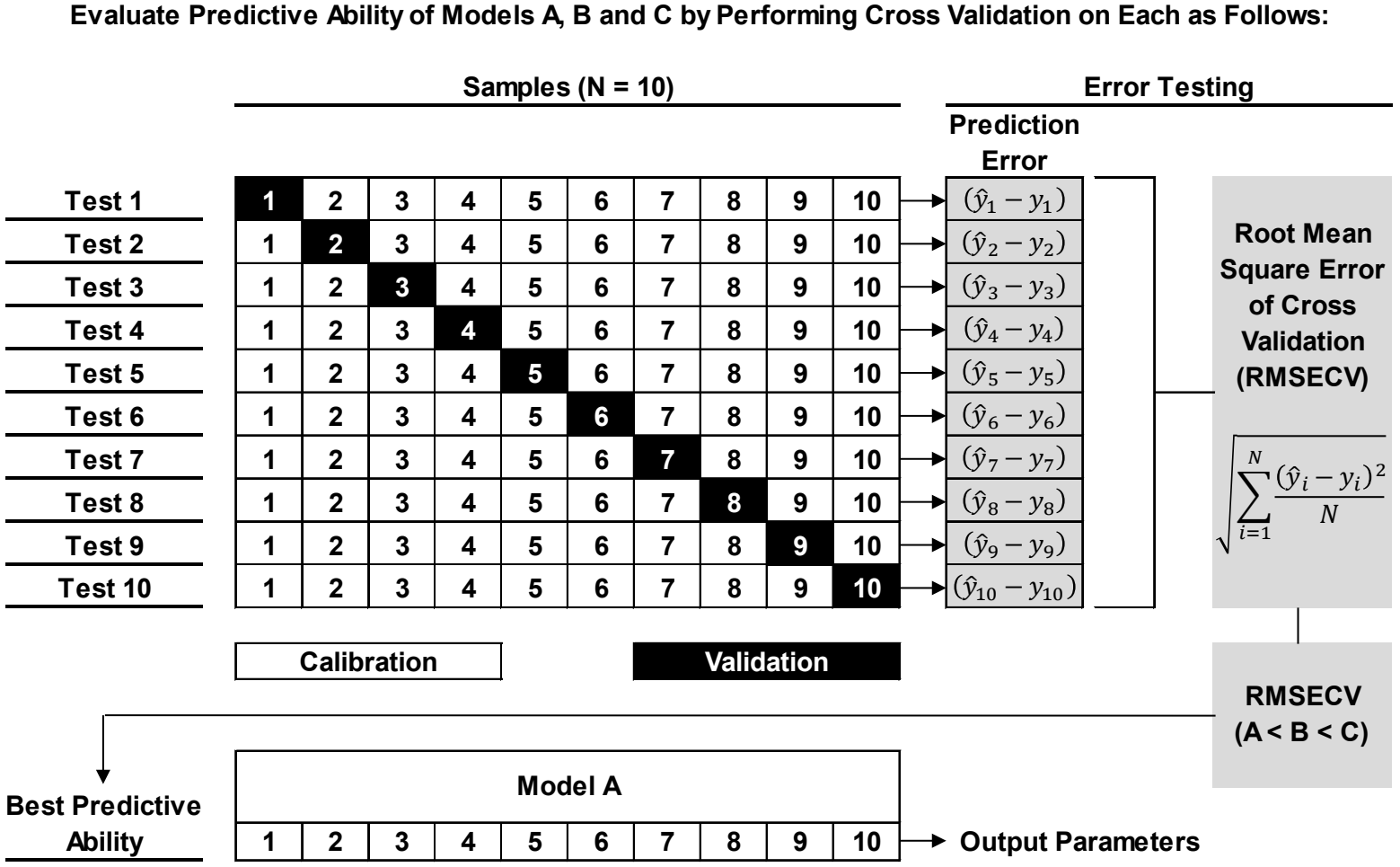


Figure 19. Cross validation of models created from a set of ten samples. [Adapted from Williams (2001) and Naes et al. (2002)].

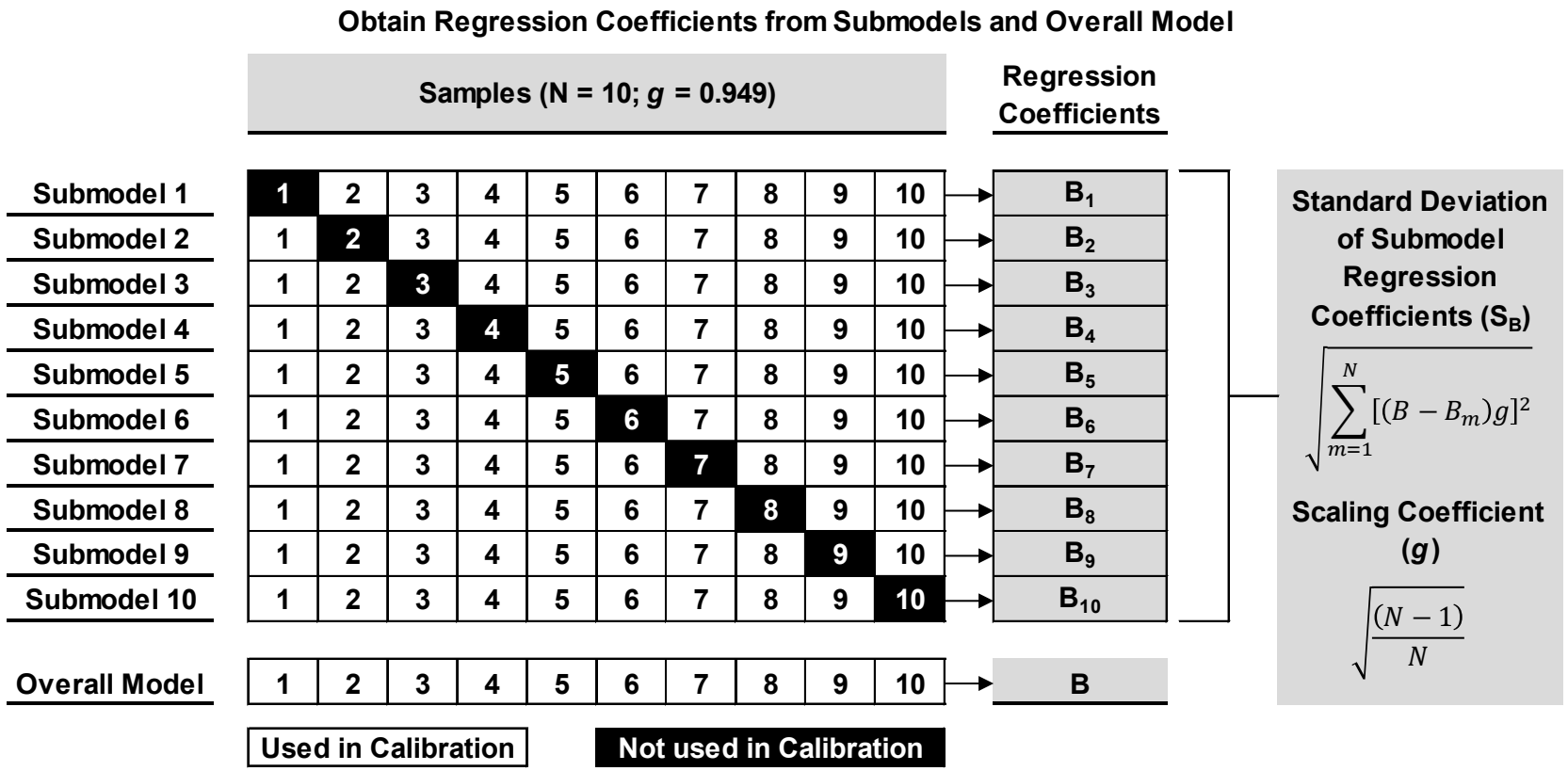
Although the "leave one out" approach at the heart of the cross validation appears again in the cross validation regression coefficients t-test (Figure 20), the two procedures differ thereafter and are used for distinct purposes. While the goal of cross validation is to evaluate model performance, that of the cross validation regression coefficients t-test is to evaluate wavenumbers to include in a model. In the former procedure, each sample is excluded in turn to test the accuracy of the model. In the latter, each sample is excluded to generate a "submodel" from the remaining samples. Like the overall model, each submodel contains a regression coefficient for each wavenumber.

Wavenumbers with larger regression coefficients in the overall model (B), whether positive or negative, are more relevant to NIR determination of the analyte of interest. Wavenumbers with a smaller standard deviation among the regression coefficients of the submodels (S_B) provide more stable results. Wavenumbers which are both relevant and stable are sought to include in good models, and can be determined by the following t-test when N samples are used [NIRCal 5.4 Manual; Martens &Martens (2000)]:

$$t = \frac{|B|\left(\sqrt{N}\right)}{(S_B)}$$

Regression coefficients can be tested for significance with the use of a t-distribution table. Those wavenumbers lacking regression coefficients statistically different from zero at the desired confidence level can be excluded from the model. Alternatively, values of t can be ranked and models created which retain wavenumbers with t values falling at or above a certain cutoff limit.

In the initial description of the cross validation regression coefficients t-test, Martens and Martens (2000) used the equation for S_B reproduced in Figure 20. This formula used a scaling coefficient g which reflected terms originally proposed by Tukey (1958) for use in "leave one out" scenarios. Tukey's approach was evaluated for effectiveness and subsequently advocated by Efron (1982).



Perform t-test at Each Wavenumber using Regression Coefficient of Overall Model (B), Number of Samples (N) and Standard Deviation of Submodel Regression Coefficients (S_B)

$$t = \frac{|B|\sqrt{N}}{S_B}$$

Figure 20. Cross validation regression coefficients t-test for a model created from a set of ten samples. [Adapted from Martens and Martens (2000), Tukey (1958) and NIRCal 5.4 Software Manual].

4.4.5.4.2.3 Model Improvement by Wavenumber Selection

Values of e or t can be used to select wavenumbers to improve model quality. Like the coefficient e, the cross validation regression coefficients t-test assesses the importance (numerator) and stability (denominator) of measurements at each wavenumber to isolate the best variables. Unlike e, which tracks the largest overall changes among sample groups, t follows the most relevant wavenumbers for the analyte of interest.

It has been demonstrated that model quality can be improved significantly in an iterative process which at each step retains data only from wavenumbers with regression coefficients that pass significance testing (Martens & Martens, 2000; Westad & Martens, 2000). As a parameter of the model, regression coefficients change when areas of the spectrum found to be of little value are dropped, thus the regression coefficients t-test can be used iteratively to improve the quality of models. The value of e is a function of spectral intensities rather than an output of the model. It is thus constant for each wavenumber and cannot be used iteratively per se.

Initial models of conjugated dienes and peroxide values created with combined data sets of 7.5% (w/w) pecan oil : white rice flour samples and 7.5% (w/w) canola oil : white rice flour samples were subjected to model improvement techniques using wavenumber selection based on e and the cross validation regression coefficients t-test. Ranking of wavenumbers by the application of cutoffs based on e and the numerator of e were evaluated. The latter metric did not account for the stability among scans of the same sample and was used to determine if important wavenumbers were discarded when using the e coefficient because of variation. Significance of regression coefficient t-test values at 99.5% confidence as well as ranking of wavenumbers by the application of cutoffs based on t were also evaluated. The best models resulting from application of these improvement procedures to the combined sample set were then applied to sample sets made with either pecan or canola oil.

5. RESULTS & DISCUSSION

5.1 Preliminary Experiments using Near Infrared Spectroscopy

5.1.1 Discrimination of Fat or Oil Type in Lipid : White Rice Flour Mixtures

Eight different oils and fats were coated on white rice flour at 15% (w/w) to assess the ability of NIR to distinguish between lipids with different fatty acid composition. As shown in Figure 21, seven of these could be resolved by NIR, but safflower oil could not be distinguished from sunflower oil due to the similarity of their respective fatty acid compositions. Hence, safflower oil was omitted from the model calculated for these samples. An acceptable qualitative model (Q-Value > 0.8) was found via principal component analysis which could discriminate among samples of the seven test lipids (Figure 21 and Table 9). The main penalty assessed arose from the proximity of the clusters corresponding to different oils and fats, which indicates the degree of similarity among the samples assayed. A second penalty arose from the moderately uneven spread of samples of the same type within their respective clusters.

The property interference penalty showed a certain degree of similarity among the samples. This was to be expected considering that the composition of each sample was nominally 85% identical (white rice flour), and the remaining 15% varied only with the difference in composition among the lipids. The scores plot indicated that saturated fats (lard and palm oil) were readily distinguishable from oils with higher linoleic acid content (sunflower [and thus safflower], canola and soybean oil) along the third principal component axis. This distinction was also evident in the in the second derivative pretreated spectra from 5760 – 5900 cm^{-1} (Figure 22). This region includes wavenumbers corresponding to the first overtones of C-H stretching vibrations for methylene (5797 cm^{-1}) and methyl groups (5865 cm^{-1}). Samples of the saturated fats showed greater intensity for the methylene group signal than those of oils.

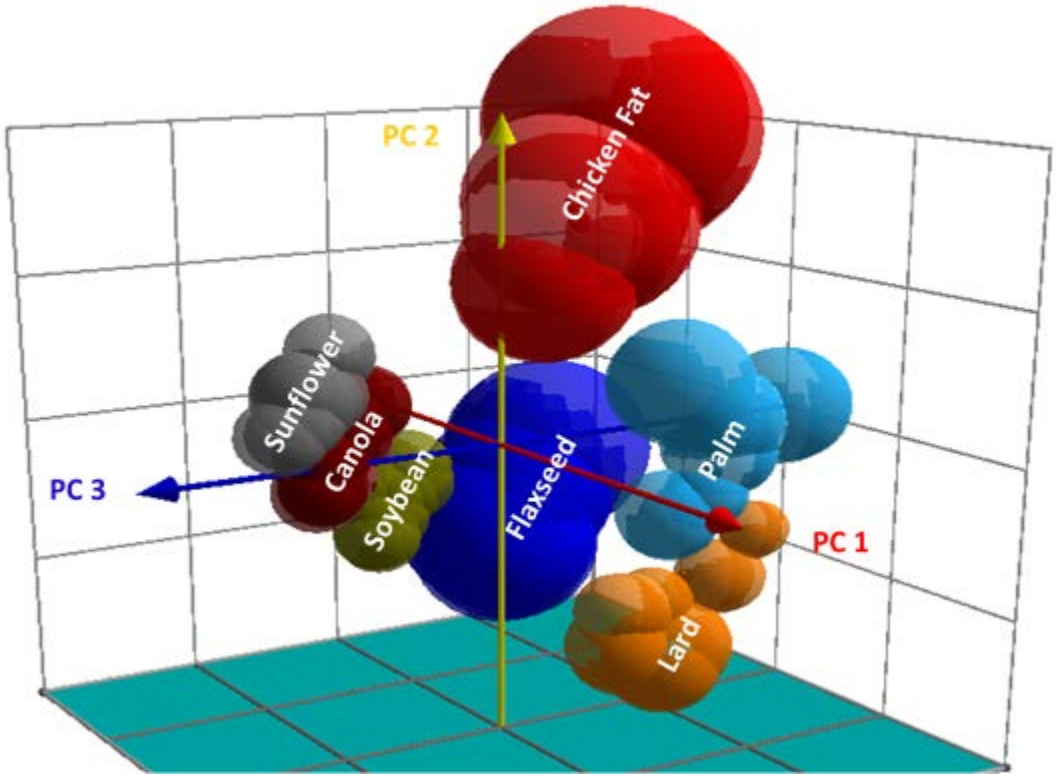


Figure 21. Scores plot of qualitative model for discrimination of 15% lipid : white rice flour mixtures.

Table 9. Q-value criteria of qualitative model for discrimination of 15% lipid : white rice flour mixtures.

Penalty	Value	Weight
C-Set False Identified (Calibration Sample in Wrong Cluster)	0	10
C-Set Not Identified (Calibration Sample Outside All Clusters)	0	10
V-Set False Identified (Validation Sample in Wrong Cluster)	0	5
V-Set False Identified (Validation Sample Outside All Clusters)	0	1
Cluster Index (Samples of Same Type Should be in Single Cluster)	0	1
Property Uniformity (Even Spread of Samples Within Clusters)	0.026112	1
Property Interference (Independence of Clusters from Each Other)	1.5385	0.1
Q-Value	**0.847485**	

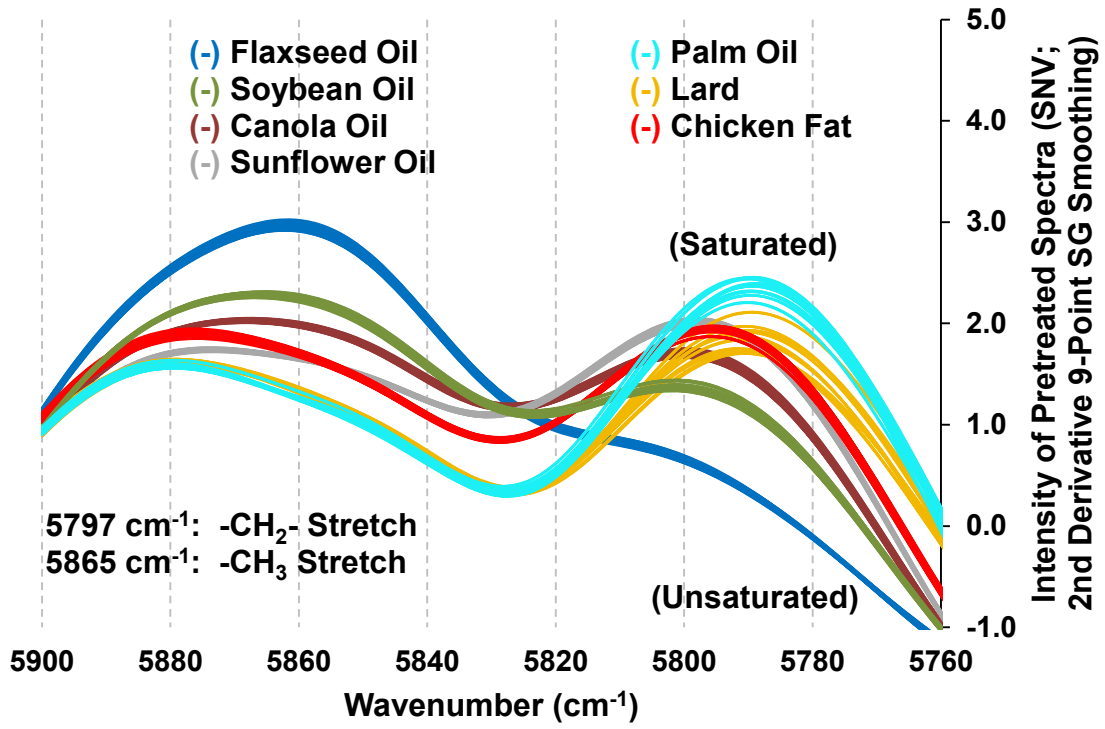


Figure 22. Pretreated spectra of different fats or oils in 15% lipid : white rice flour mixtures from 5760-5900 cm^{-1}.

The property uniformity penalty indicated a certain degree of variation in the spread of nominally identical samples within their respective clusters. Two effects related to such variation were evident. First, samples of the saturated fats displayed greater variability in scores plots than those of oils (Figure 21). This is believed to be an artifact of the sample preparation process. Small pieces of saturated fat samples often stuck to the die upon removal of the sample tablet, rendering the surface scanned by NIR uneven. Since this experiment involved a single point measurement directed at the center of the tablet's surface, the effect of physical scattering was more pronounced for these samples in comparison to the oil samples which ejected cleanly and had a uniform flat surface. This scattering was particularly evident for lard and palm oil and is apparent in the 5760 – 5820 cm^{-1} region of the second derivative spectra (Figure 22).

The second effect was a noticeable streaking of scores for any given sample. For all samples, clusters extend from the lower left to the upper right (Figure 21). Subsequent analyses of mixtures with nominal oil content of 10% and 7.5% indicated that the streaking was due to inhomogeneity as scores of samples with different amounts of oil fell along the same axis of variation. Each measurement in Figure 21 was a single point reading of a non-homogeneous mixture. Accordingly, certain samples contained higher and lower oil content than the nominal 15%. This effect was also seen in numerous other experiments.

In conclusion, NIR was determined capable of distinguishing among different oils and fats when present in a mixed solid food model system at a nominal level of 15% by weight. NIR analyses were acceptable in the qualitative context for mere sample identification, but sample presentation issues, particularly for saturated fats, as well as sample inhomogeneity gave rise to confounding effects that might significantly impair the fit of quantitative models. These considerations must be addressed if NIR is to be used for quantitative analyses of lipid oxidation.

5.1.2 Discrimination of Oxidation Status in Lipid : White Rice Flour Mixtures

Having demonstrated that NIR could distinguish between different oils in food model systems, the next step was to determine whether extents of lipid oxidation could be differentiated by NIR and which products could be identified. Accordingly, 15% (w/w) oil: white rice flour mixtures were prepared with soybean, sunflower, safflower or canola oils and incubated at 60°C for up to three weeks. Results of qualitative NIR models from principal component analysis (PCA) were mixed. Models for sunflower and safflower oils were not attainable under the conditions investigated. A model for soybean oil could discriminate only among the first three time points (T0, one and two weeks incubation). The scores plot for this model is shown in Figure 23. Q-value criteria for this model are shown in Table 10. An acceptable qualitative model that discriminated among all four time points (T0 and one, two or three weeks of

incubation) was only obtained for canola oil. The scores plot for this model is shown in Figure 24. Q-value criteria for this model are shown in Table 11.

To produce an acceptable qualitative model, PCA requires the ability to differentiate between each time point. If any two cannot be discriminated, the model will fail. Given the similarity of the sunflower and safflower oils tested as noted in the experiment above, as well as the high polyunsaturated fatty acid content of sunflower, safflower and soybean oils (67.5, 77.7 and 53.7% linoleic acid (18:2), respectively), the inability to distinguish samples at different time points likely reflected the difference in times necessary to degrade these oils. Among those tested, the canola sample was apparently the most stable. As all four time points of the canola incubation were discernible, the other oil samples reached a state of degradation by the third week or sooner which was undiscernible from at least one other time point. This is substantiated by an apparent circular pattern through which samples move during oxidation as shown in scores plots of the soybean and canola models (Figure 23 and Figure 24, respectively).

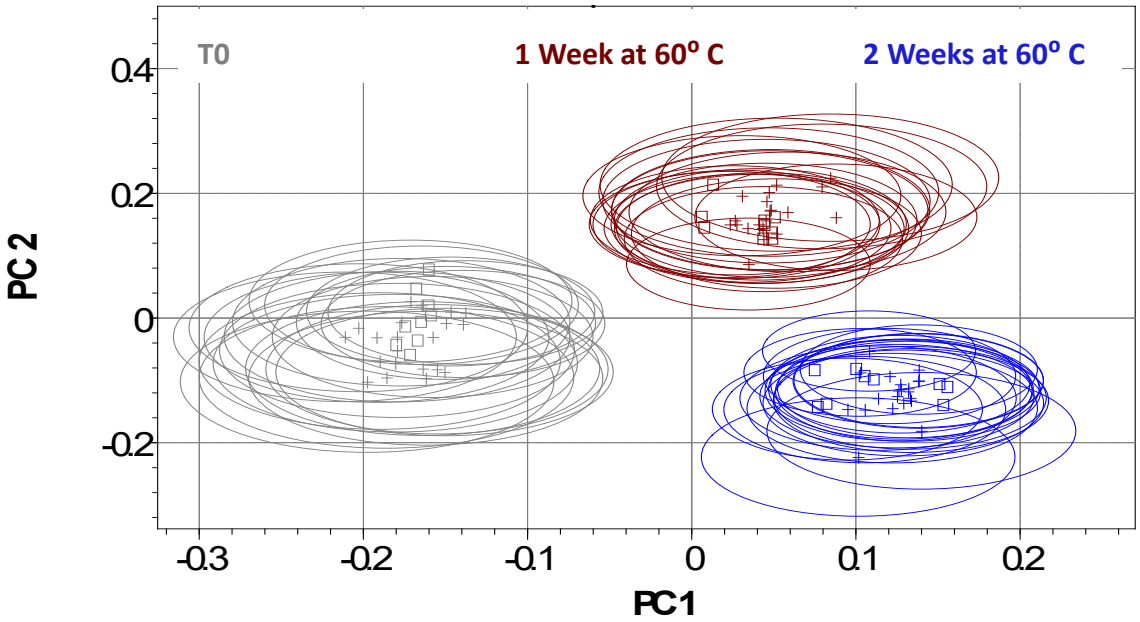


Figure 23. Scores plot of discrimination of oxidation status of soybean oil in 15% soybean oil : white rice flour mixtures. [Due to data loss, image shown from NIRCal 5.4 with tolerance radii (ellipses) about scores of calibration (+) and validation (□) samples.]

Table 10. Q-value criteria of qualitative model for discrimination of oxidation status of soybean oil in 15% soybean oil : white rice flour mixtures.

Penalty	Value	Weight
C-Set False Identified (Calibration Sample in Wrong Cluster)	0	10
C-Set Not Identified (Calibration Sample Outside All Clusters)	0	10
V-Set False Identified (Validation Sample in Wrong Cluster)	0	5
V-Set False Identified (Validation Sample Outside All Clusters)	0	1
Cluster Index (Samples of Same Type Should be in Single Cluster)	0	1
Property Uniformity (Even Spread of Samples Within Clusters)	0.0158566	1
Property Interference (Independence of Clusters from Each Other)	0.903865	0.1
Q-Value	**0.90396**	

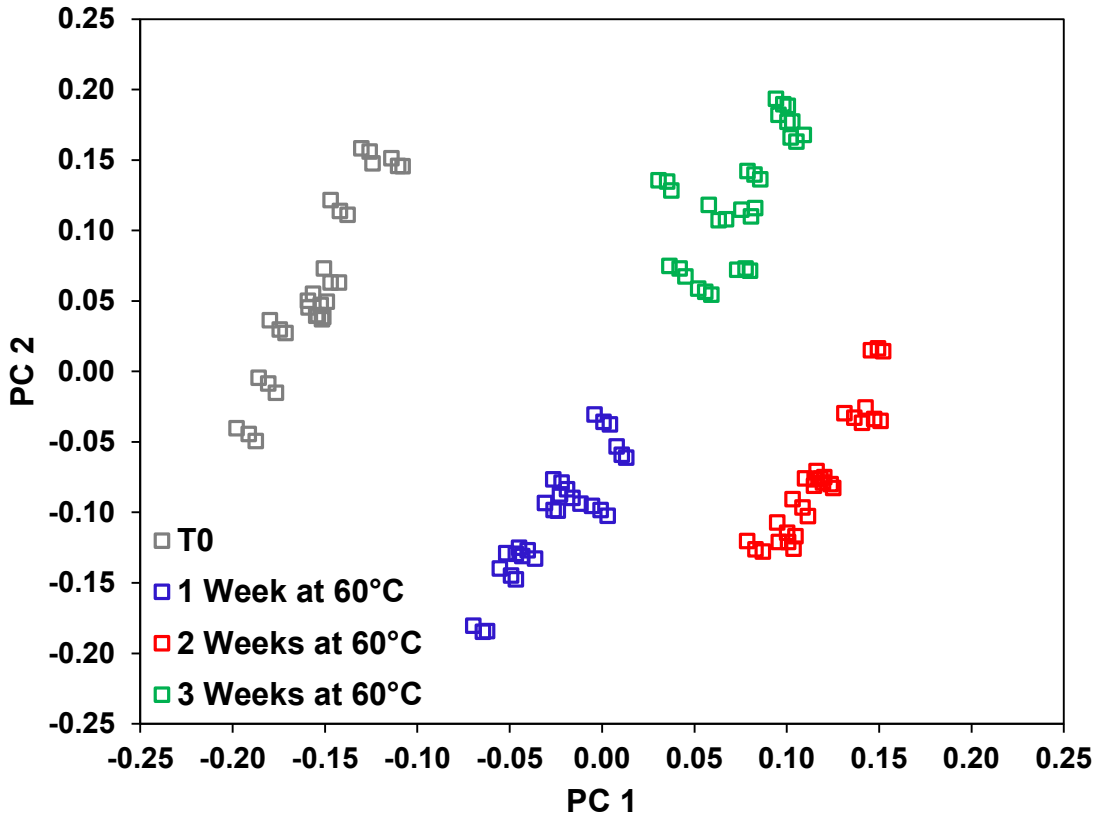


Figure 24. Scores plot of discrimination of oxidation status of canola oil in 15% canola oil : white rice flour mixtures. Each square represents the score of an individual spectrum acquired on the sample. Three replicate scans were made on each of ten tablets pressed from each sample.

Table 11. Q-value criteria of qualitative model for discrimination of oxidation status of canola oil in 15% canola oil : white rice flour mixtures.

Penalty	Value	Weight
C-Set False Identified (Calibration Sample in Wrong Cluster)	0	10
C-Set Not Identified (Calibration Sample Outside All Clusters)	0	10
V-Set False Identified (Validation Sample in Wrong Cluster)	0	5
V-Set False Identified (Validation Sample Outside All Clusters)	0	1
Cluster Index (Samples of Same Type Should be in Single Cluster)	0	1
Property Uniformity (Even Spread of Samples Within Clusters)	0.02505	1
Property Interference (Independence of Clusters from Each Other)	1.634	0.1
Q-Value	**0.8414**	

In both instances, separation of the samples was possible on two principal component axes. Interestingly, scores in the canola model appeared to follow a circular path as oxidation progressed. This was likely in keeping with the cyclic nature of lipid oxidation frequently mentioned in the literature. Concentrations of early products, conjugated dienes and peroxides, rise to a certain level before falling as they are converted to secondary product carbonyls. The soybean model scores also appeared to follow such a path, though they were less conclusive on this point as soybean oil oxidized earlier, rendering the 3 week sample indistinguishable from one of the earlier samples and causing the four point model to fail.

Incubation occurred in bottles open to air, which would have allowed volatile aldehydes to evaporate from the samples. Although it is noted by some authors (Dellarosa et al., 2015; Yildiz et al., 2001) that signals for the direct detection of carbonyls are lacking in the NIR region anyhow, the loss of volatile aldehydes from late stage oxidation samples would further render their spectra similar to early stage samples with equivalent levels of primary oxidation products.

It was also noteworthy that the canola model was obtained from PCA after selecting certain wavenumbers from the full range of 4000-10,000 cm^{-1}. This model included spectral ranges from 4152-4216 cm^{-1}, 4300-4552 cm^{-1}, 4848-5100 cm^{-1}, 5148-5480 cm^{-1}, 6980-7340 cm^{-1} and 7500-7652 cm^{-1}. Loadings for both principal component axes are shown in Figure 25. The importance of wavenumbers around 5000-5200 cm^{-1} as well as at 7000-7200 cm^{-1} was consistently evident in preliminary oxidation studies of different oils on white rice flour.

Possible chemical assignments for important wavenumbers from each of the loading spectra are shown in Table 12. It is apparent therein that wavenumbers consistent with those determined from previous lipid oxidation studies (Appendix A) as well as those relevant to the flour matrix (starch, protein) and possible interferents (moisture) are represented. This underscores the complexity of the problem as well as the need to account for sample handling and presentation issues as fully as possible for fine analyses using NIR.

As was the case with the model from the lipid discrimination experiment above, streaking of scores among nominally identical samples was observed in the canola oxidation model. The scores for replicate scans of the ten NIR sample tablets made from canola oil at T0 are shown in Figure 26a, with the tablet number closest to the first replicate and subsequent replicates connected in order by a line for that tablet. Although scores of each tablet fell randomly within the cluster along both principal component axes, replicate scans for a given tablet drifted systematically, primarily along the axis for principal component 1. These effects were also evident in scores of replicates at each incubation time point (Figure 26b-d). The random scatter was likely due to inhomogeneity among the samples. The drifting was likely due to sample temperature or moisture content as neither was controlled and either could result in a systematic change with time. This is discussed further in Section 5.2.4.2.

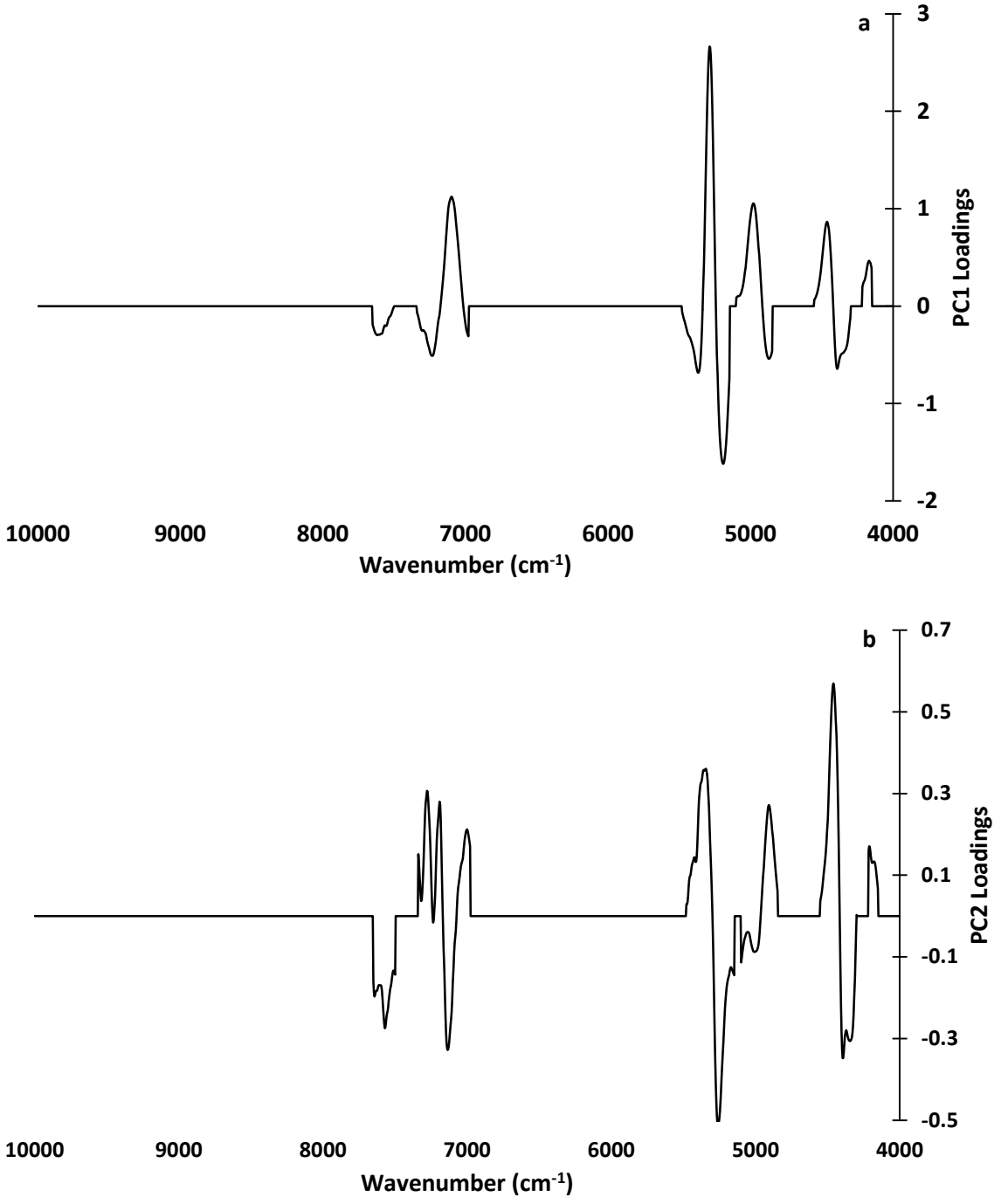


Figure 25. Loadings plots for 15% canola oxidation status model: (a) Principal component 1; (b) Principal component 2.

Table 12. Possible chemical assignments from 15% canola oil oxidation status model loadings: (a) Principal component 1; (b) Principal component 2.

a. Principal component 1

Rank	Wavenumber (cm^{-1})	Loading	Possible Assignment(s)
1	5284	2.6567	**Starch:** 5263*; **Carboxylic Acid:** 5263*, 5263 (Holman & Edmondson, 1956); **Hydroxyl:** 5241*
2	5192	-1.6196	**Amide**: 5208*; **Water:** 5165 (Realini et al., 2004), 5181 (Berzaghi et al., 2005), 5187 (Cozzolino et al., 2005), 5195 (Kaddour et al., 2006) and 5208 (Karlsdottir et al., 2014); **Possible Secondary Oxidation Product**: 5219 (Takamura et al., 1995)
3	7096	1.1237	**Hydroxyl:** 7092*, 7042 (Holman et al., 1958); **Hydroperoxides:** 7068 (Wojcicki et al., 2015)
4	4980	1.0551	**Hydroxyl:** 4975 (Holman et al., 1958); **Carbonyl from Liberated Fatty Acids:** 4980 (Cho et al., 1998); **Starch:** 5000*; **Amide**: 5000*
5	4464	0.8645	**Possible Secondary Oxidation Product:** 4456 (Takamura et al., 1995); **Amino Acid:** 4460*; **Starch:** 4440*; **Terminal Epoxides**: 4532 (Peck et al., 1987), 4545 (Goddu & Delker, 1958)
6	5368	-0.6839	**Polymer Content of Sunflower Oil:** 5400 (El-Rafey et al., 1988)
7	4392	-0.6452	**Starch:** 4394*; **Methyl**: 4386 (Holman & Edmondson, 1956)
8	4872	-0.5409	**Linoleic Acid Content of Sunflower Oil:** 4873 (El-Rafey et al., 1988); **Protein:** 4864 (Berzaghi et al., 2005), 4878*; **Hydroperoxides:** 4831 (Holman et al., 1958)
9	7232	-0.5107	**Polar Content of Sunflower Oil:** 7267 (El-Rafey et al., 1988)
10	4168	0.4646	-
11	7620	-0.2968	-

* Indicates Data from NIRCal 5.4 Software Chemical Bonding Module

Table 12. (Continued) Possible chemical assignments from 15% canola oil oxidation status model loadings: (a) Principal component 1; (b) Principal component 2.

b. Principal component 2

Rank	Wavenumber (cm^{-1})	Loading	Possible Assignment(s)
1	4460	0.5696	**Possible Secondary Oxidation Product:** 4456 (Takamura et al., 1995); **Amino Acid:** 4460*; **Starch:** 4440*; **Terminal Epoxides**: 4532 (Peck et al., 1987), 4545 (Goddu & Delker, 1958)
2	5260	-0.5131	**Starch:** 5263*; **Carboxylic Acid:** 5263*, 5263 (Holman & Edmondson, 1956); **Hydroxyl:** 5241*
3	5344	0.3602	**Polymer Content of Sunflower Oil:** 5400 (El-Rafey et al., 1988)
4	5360	0.3575	**Polymer Content of Sunflower Oil:** 5400 (El-Rafey et al., 1988)
5	4392	-0.3479	**Starch:** 4394*; **Methyl**: 4386 (Holman & Edmondson, 1956)
6	7136	-0.3280	**Hydroxyl:** 7143 (Cozzolino et al., 2005)
7	7280	0.3063	**Polar Content of Sunflower Oil:** 7267 (El-Rafey et al., 1988)
8	4344	-0.3059	**Methylene:** 4348 (Holman & Edmondson, 1956), 4348 (Kaddour et al., 2006); **Conjugated Systems:** 4348 (Holman & Edmondson, 1956)
9	7192	0.2800	**Methylene:** 7168*; **Polymer Content of Sunflower Oil:** 7163 (El-Rafey et al., 1988); **Hydroxyl:** 7143 (Cozzolino et al., 2005)
10	7572	-0.2745	-
11	4908	0.2719	**Amide:** 4926*

* Indicates Data from NIRCal 5.4 Software Chemical Bonding Module

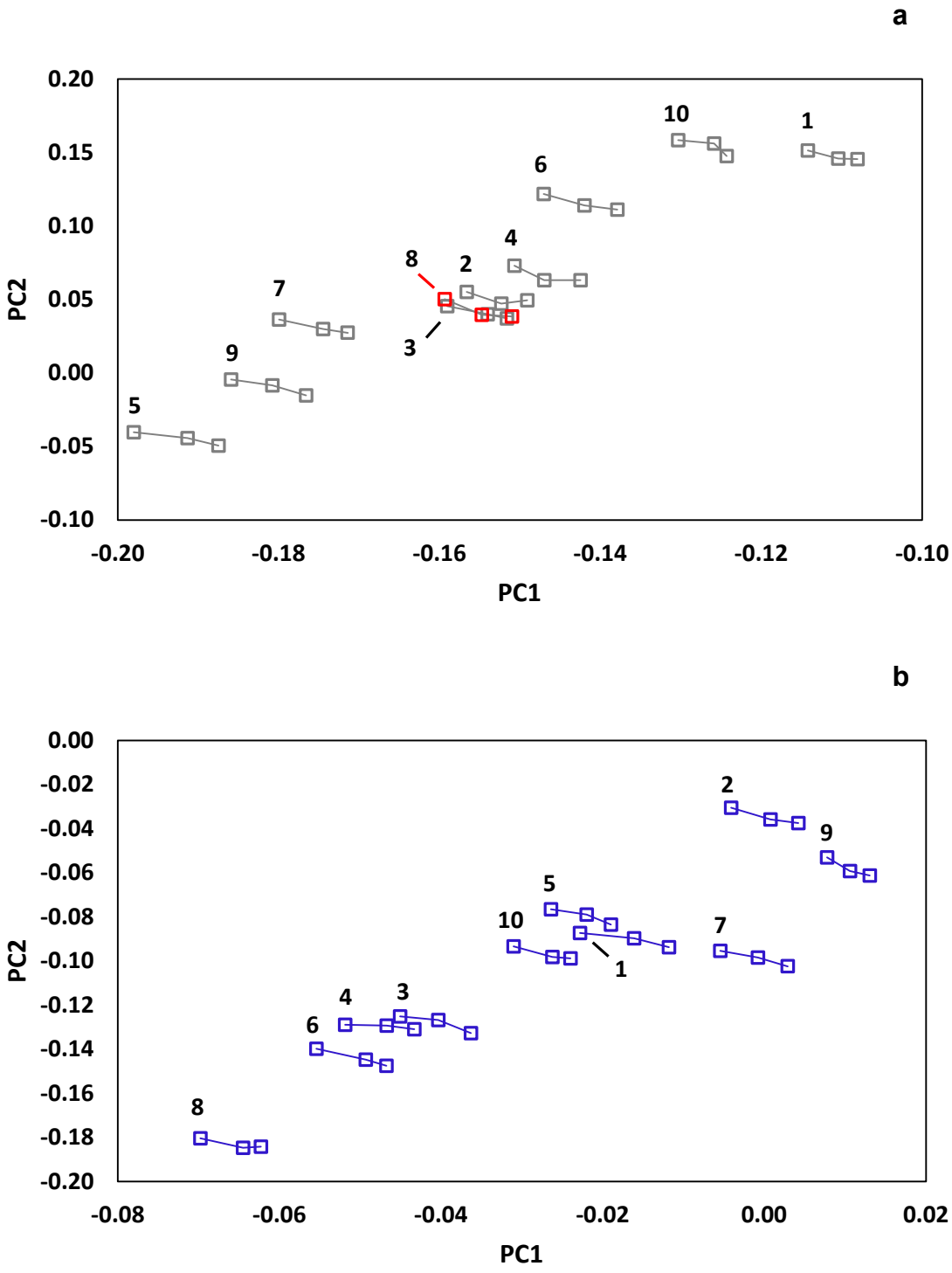


Figure 26. 15% Canola oil oxidation status model scores: (a) T0; (b) 1 Week; (c) 2 Weeks; and (d) 3 Weeks incubation. Each number indicates the first replicate of each sample tablet in the order assayed. Replicate scans are connected by lines in the order acquired.

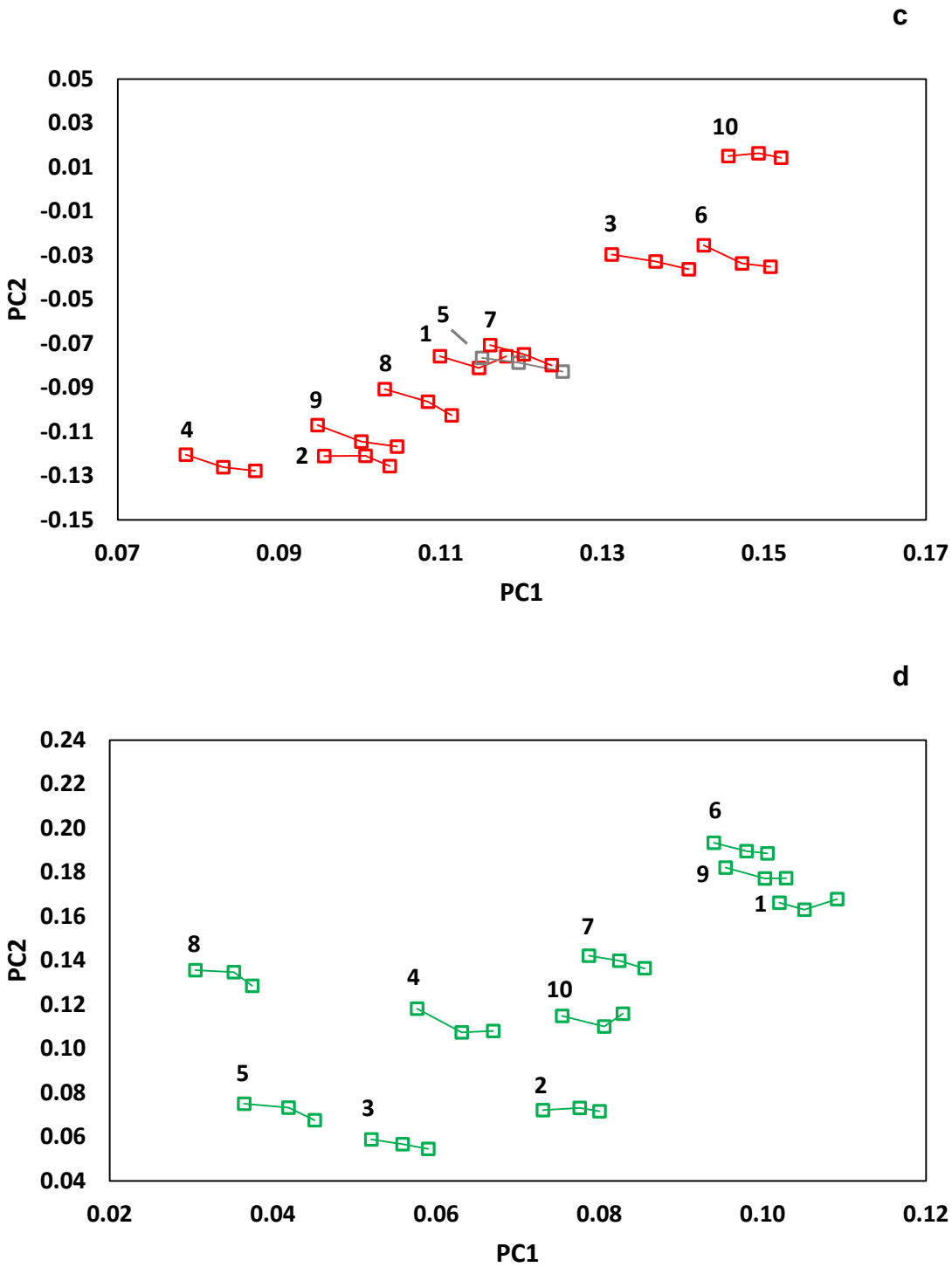


Figure 26. (Continued) 15% Canola oil oxidation status model scores: (a) T0; (b) 1 Week; (c) 2 Weeks; and (d) 3 Weeks. Each number indicates the first replicate of each sample tablet in the order assayed. Replicate scans are connected by lines in the order acquired.

5.1.3 Effect of Sample Packing on Variability of NIR Response

5.1.3.1 15% Canola Oil: White Rice Flour Samples

Effects of sample packing on NIR spectra were assessed using tablets of 15% (w:w) canola oil: white rice flour samples pressed uniformly at four discrete pressures using the lever press. FT-NIR was unable to fully discriminate among nominally identical samples compacted at different forming pressures from 340 to 415 psi, as shown in the overlap of clusters in the scores plot (Figure 27). However, when the highest pressure samples at 415 psi were excluded from the analysis, FT-NIR was able to resolve the same samples into three distinct groups according to forming pressure and thus sample packing (Figure 28).

Sample inhomogeneity was the likely explanation for these results. Pressed samples were scanned as single point measurements without rotation. Increased forming pressure caused a decrease in scores along PC2 (Figure 27 and Figure 28). The white rice flour matrix was relatively hydrophilic, and the application of too high a forming pressure caused the oil to separate from the matrix, increasing inhomogeneity and confounding the results. This would explain why the 415 psi cluster was unable to be resolved from the 390 psi cluster. It would also explain why there was no overlap of the 415 psi samples with the 340 psi samples and only minimal overlap at the edge of the 365 psi cluster in Figure 27.

Although the application of a forming pressure was devised primarily to minimize variation from scattering effects in the solid matrix, early experiments also indicated practical effects for sample handling. Pressed samples of white rice flour mixed with only minor amounts of oil were extremely friable. Elevating oil concentrations and increasing forming pressures both reduced breakage of sample tablets during transfer from the forming die to the FT-NIR sample holder. Based upon the results of this experiment, it was expected that the pressure which induced inhomogeneity would vary inversely with the oil content of the sample.

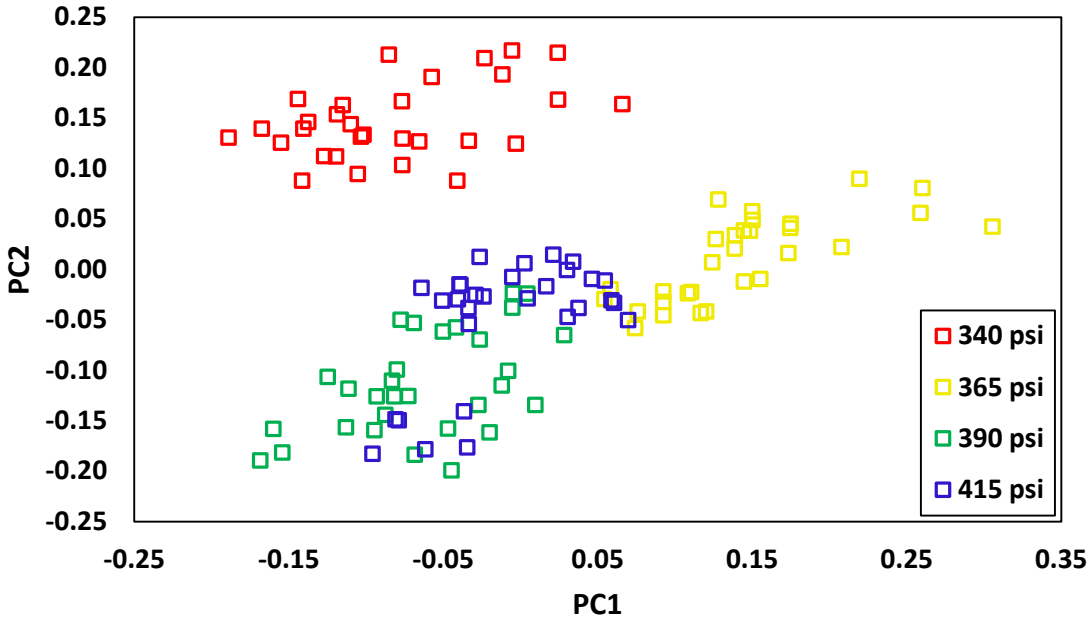


Figure 27. Scores plot showing effects of pressure on variability in 15% canola oil : white rice flour samples using all four forming pressures assayed.

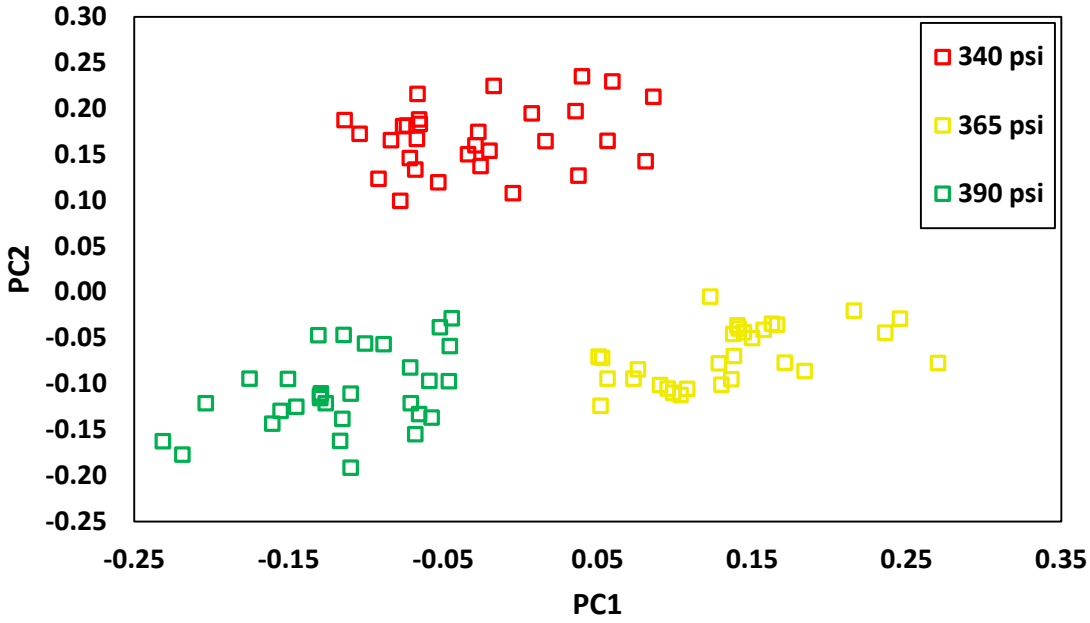


Figure 28. Scores plot showing effects of pressure on variability in 15% canola oil : white rice flour samples using the lowest three forming pressures assayed.

A review of numerous cereal-based food products at the inception of our work indicated most had oil contents of less than 10%. Our initial experiments were performed with 15% oil to ensure that the alterations of interest solely in this minor constituent could be detected by FT-NIR. Given the inverse relationship between forming pressure and oil content as well as other preliminary results indicating the sensitivity of FT-NIR, we undertook an investigation of samples incorporating only 7.5% oil.

5.1.3.2 7.5% Canola Oil: White Rice Flour Samples

Results of the experiment to assess the ability of NIR to discriminate among 7.5% (w:w) canola oil: white rice flour sample tablets pressed uniformly at different pressures are shown in the scores plot in Figure 29 and Q-Value criteria in Table 13. In this experiment, NIR was able to discriminate among samples of identical nominal composition subjected to pressures differing by only 25 psi. Accordingly, the use of NIR to evaluate forming pressures in powdered samples with modest oil composition appeared effective, in keeping with its use at much higher pressures employed in pharmaceutical tableting (Guo et al., 1999; Roggo et al., 2005).

Conversely, if identical samples were packed with variable pressures, e.g. either different loading into vials or different packing into tablets, the differences in packing would increase variation among the corresponding NIR spectra and could thereby increase scatter among scores. Thus, controlling sample packing must be a critical consideration for reducing scatter and eliminating overlap of sample clusters. It is essential for accurate quantitation of lipid oxidation and other assays tracking minor changes in samples using NIR.

Whether pressure differences confounded analysis of lipid oxidation in the initial ASL studies described above depends on the loadings of the pressure discrimination model in comparison to those of a lipid oxidation model. If there is overlap in the wavenumbers critical to the models, it is more likely that pressure differences in lipid oxidation samples would interfere with that analysis.

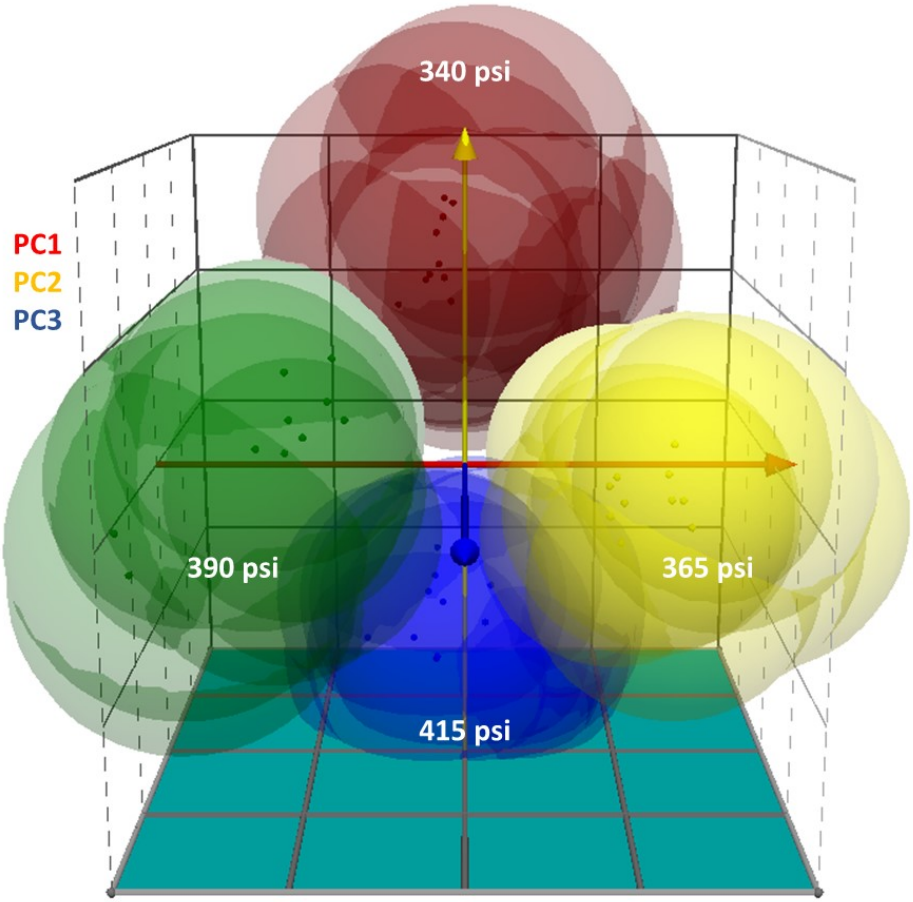


Figure 29. Scores plot of pressure discrimination model in 7.5% canola oil : white rice flour samples using all four forming pressures assayed. Spheres indicate tolerance radii about sample scores shown as points within to aid three-dimensional visualization.

Table 13. Q-value criteria of qualitative model for pressure discrimination in 7.5% canola oil : white rice flour samples.

Penalty	Value	Weight
C-Set False Identified (Calibration Sample in Wrong Cluster)	0	10
C-Set Not Identified (Calibration Sample Outside All Clusters)	0	10
V-Set False Identified (Validation Sample in Wrong Cluster)	0	5
V-Set False Identified (Validation Sample Outside All Clusters)	0	1
Cluster Index (Samples of Same Type Should be in Single Cluster)	0	1
Property Uniformity (Even Spread of Samples Within Clusters)	0.001667	1
Property Interference (Independence of Clusters from Each Other)	0.6691	0.1
Q-Value	**0.9358**	

Table 14 shows the eigenvalues for all four principal components of the pressure discrimination model, including the three principal component axes along which the samples can be discriminated as well as a fourth corresponding to errors in the data. The eigenvalue for each principal component indicates the amount of variation among samples in the model attributable to that principal component. Principal components one through three account for 38.4%, 26.0% and 24.6% of the variation in the set of samples, respectively.

Figure 30 shows the loadings spectra for the first three principal components, with red highlighted areas in each corresponding to wavenumber ranges used in the qualitative canola oxidation model above. Although the most critical wavenumbers in the pressure discrimination model are between 9000 – 10,000 cm^{-1}, there are important features for all three principal components in this pressure model falling within the 4300-4552 cm^{-1} range used in the canola oxidation model. This range is particularly important in loadings of the first and most important principal component in the pressure discrimination model (Figure 30a). Also, changes in sample packing, arising here from different forming pressures, are a bulk phenomenon. Thus, packing differences among samples likely contributed to variation in sample scores of the qualitative canola oxidation model and forming pressure should be controlled to minimize variation during quantitative analyses.

Table 14. Eigenvalues of qualitative model for pressure discrimination in 7.5% canola oil : white rice flour samples.

	Eigenvalue	% of Total Variation
Principal Component 1	0.0323	38.4
Principal Component 2	0.0219	26.0
Principal Component 3	0.0207	24.6
Principal Component 4 (Error)	0.0092	10.9
Total	**0.0841**	**100**

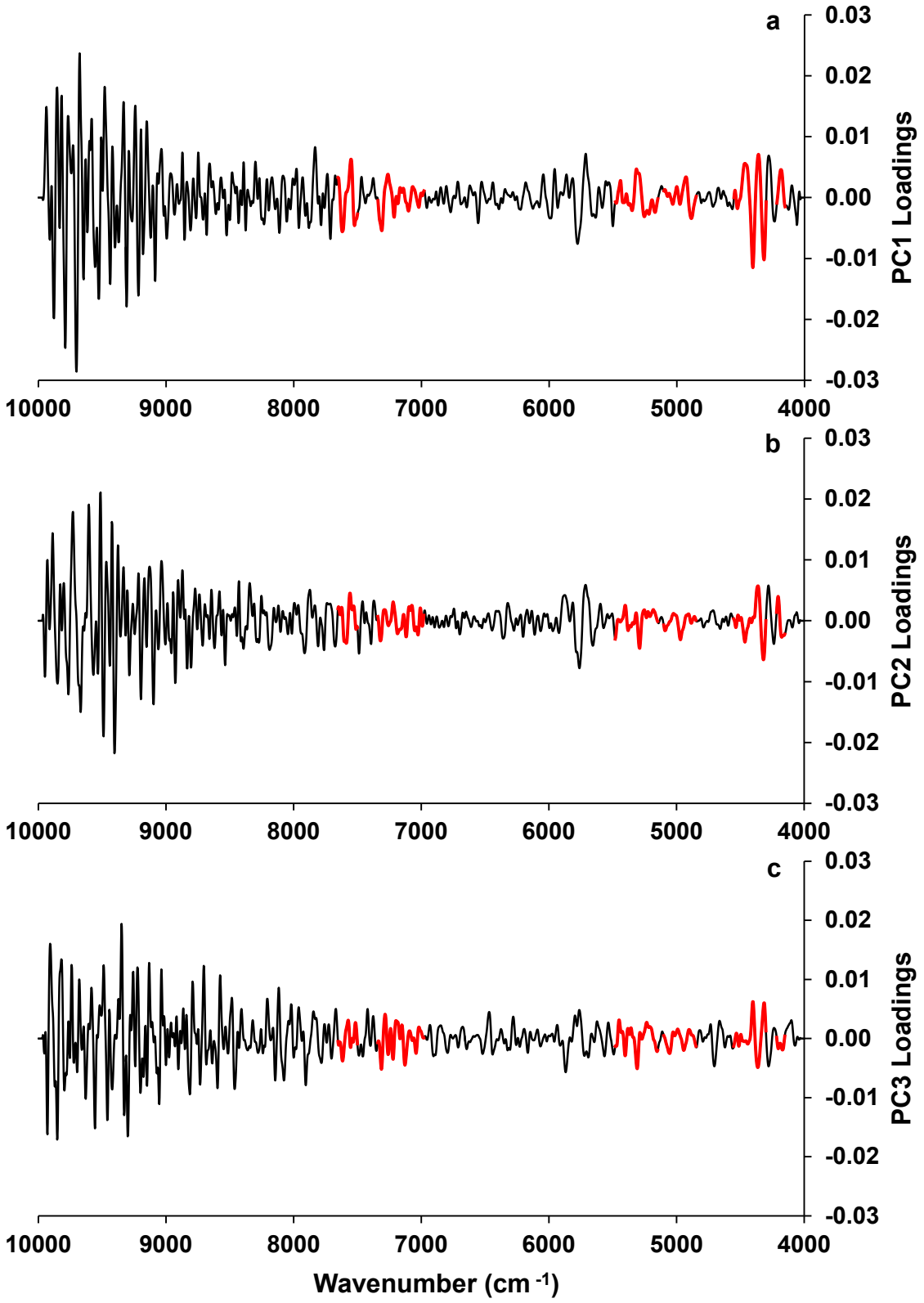


Figure 30. Loadings plots for 7.5% canola oil pressure discrimination model: (a) Principal component 1; (b) Principal component 2; (c) Principal component 3. Wavenumber ranges important in loadings of the canola oxidation model are shown in red.

5.1.4 Effect of Sample Rotation on Variability in NIR Spectra

Rotation of samples during spectral acquisition markedly decreased variation in scores plots for both single and multiple samples regardless of the forming pressure applied.

5.1.4.1 Variability Among Replicate Scans of a Single Sample

Results of replicate FT-NIR scans of a single sample of 10% (w:w) canola oil: white rice flour assayed as a powder or after tableting at 415 psi are shown in Figure 31. Both panels of Figure 31 are on the same scale and show an identical area of principal component space. Even in the 1.125 inch diameter sample assay system employed here, rotation significantly reduced the variation among scans of a single sample on both principal component axes.

Standard deviation data for the scores of 21 scans of each sample are shown in Table 15 and Figure 32. Sample rotation reduced the standard deviation of scores along Principal Component 1 by nearly 80% for both powdered and 415 psi pressed samples. Rotation also reduced the standard deviation of scores along Principal Component 2 for these samples by 42% and 86%, respectively. Although scores from stationary scans of the pressed sample were more variable than those of the powdered sample along both principal components, when samples were rotated, tableting of the sample slightly increased the standard deviation of scores along Principal Component 1 but greatly decreased the standard deviation along Principal Component 2.

Closer inspection of scores for replicates of the pressed sample showed a trend towards more positive values of Principal Component 1 from early to late samples (Figure 33). This is consistent with the trend observed in each of the samples in Figure 26. Conversely, a trend towards more positive values of Principal Component 2 is seen among replicates of the powdered sample (Figure 34). Each trend may represent an interfering effect particular to the state of the sample. A powdered sample predominantly comprised of white rice flour is expected to be more susceptible to moisture, while a pressed sample may exhibit temperature changes while cooling during scanning to dissipate heat generated by the forming process.

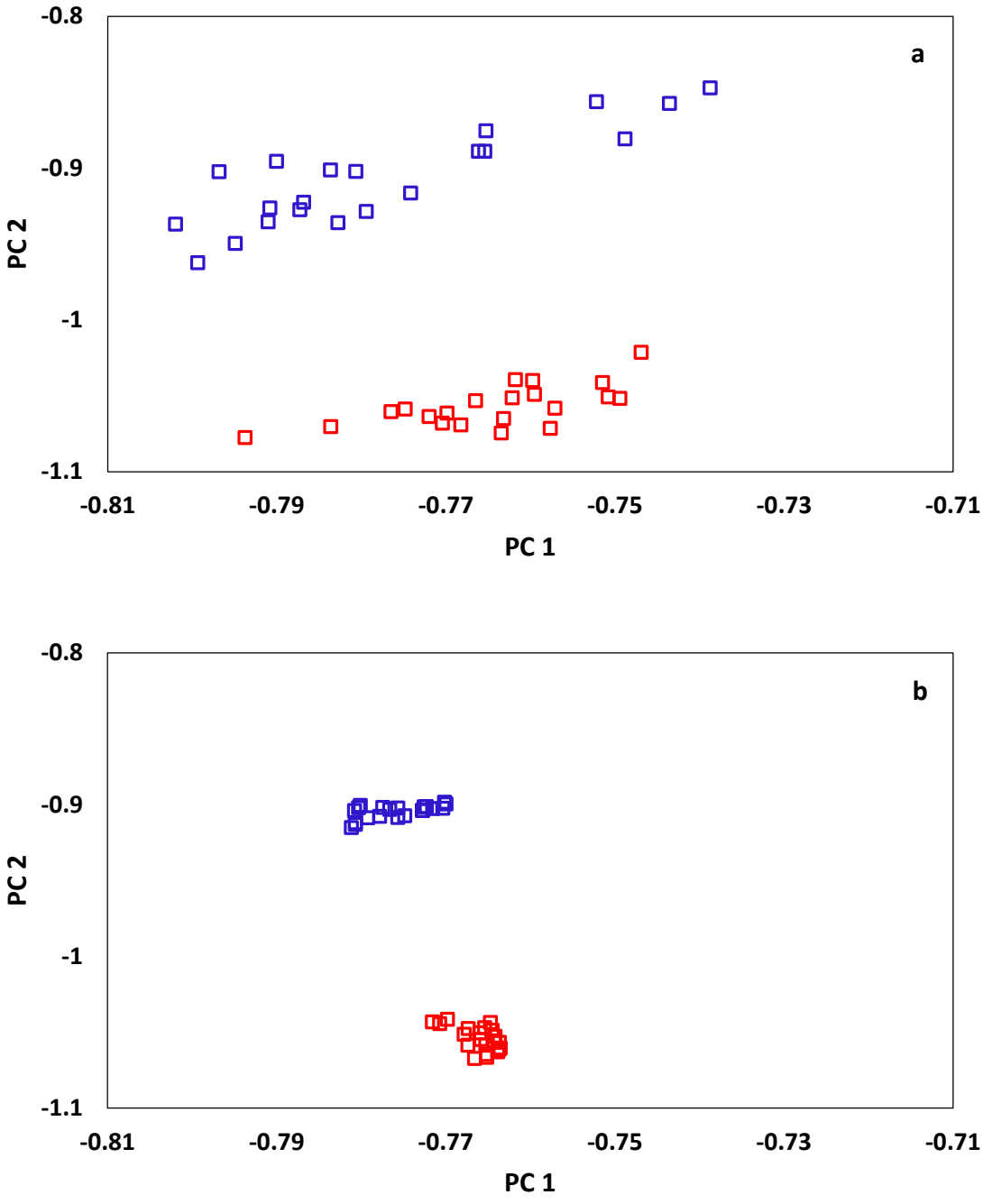


Figure 31. Scores plots of a 10% canola oil: white rice flour sample: (a) Stationary; (b) Rotating [□ Powder; □ 415 psi Tablet]. Each square indicates the score of a single replicate stationary scan made at or rotating scan starting at a distinct point of the sample.

Table 15. Standard deviation of scores of one 10% canola oil : white rice flour sample.

21 Scans Each of 1 Sample	Powder (n=21)		415 psi Tablet (n=21)	
	Stationary	Rotating	Stationary	Rotating
Principal Component 1	1.16×10^{-2}	2.33×10^{-3}	1.88×10^{-2}	3.98×10^{-3}
Principal Component 2	1.38×10^{-2}	8.00×10^{-3}	3.17×10^{-2}	4.57×10^{-3}

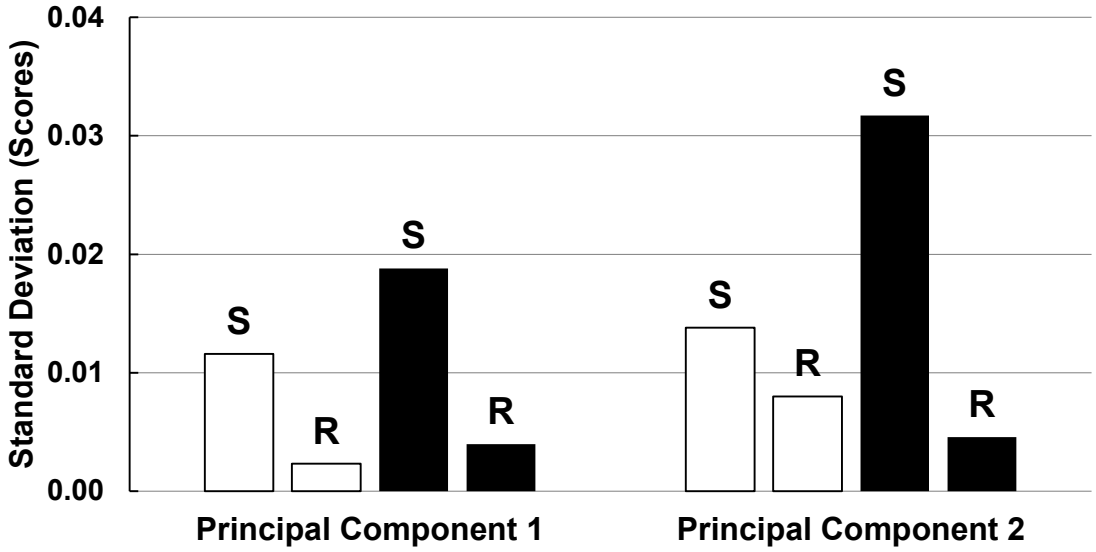


Figure 32. Standard deviation among scores of 21 scans of a single 10% canola oil : white rice flour sample [☐ Powder; ■ 415 psi Tablet; Stationary (S); Rotating (R)].

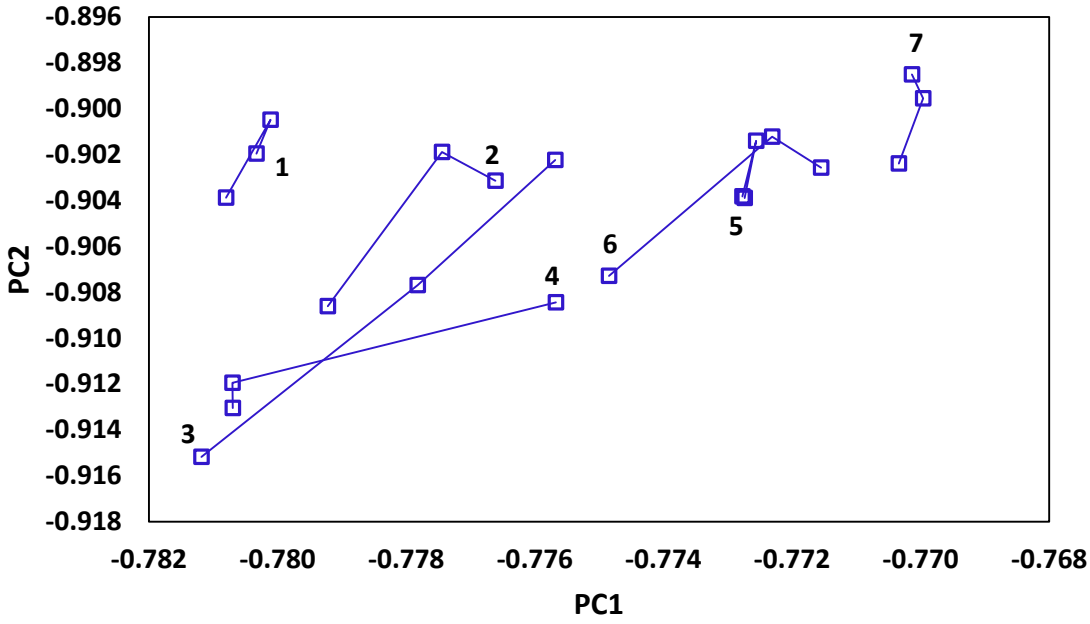


Figure 33. Expanded view of scores plot from scans of a single pressed 10% canola oil : white rice flour sample acquired with rotation. Groups of three scans acquired at 120° angles relative to each other are connected by lines. Numbers indicate the first scan of each group in temporal order.

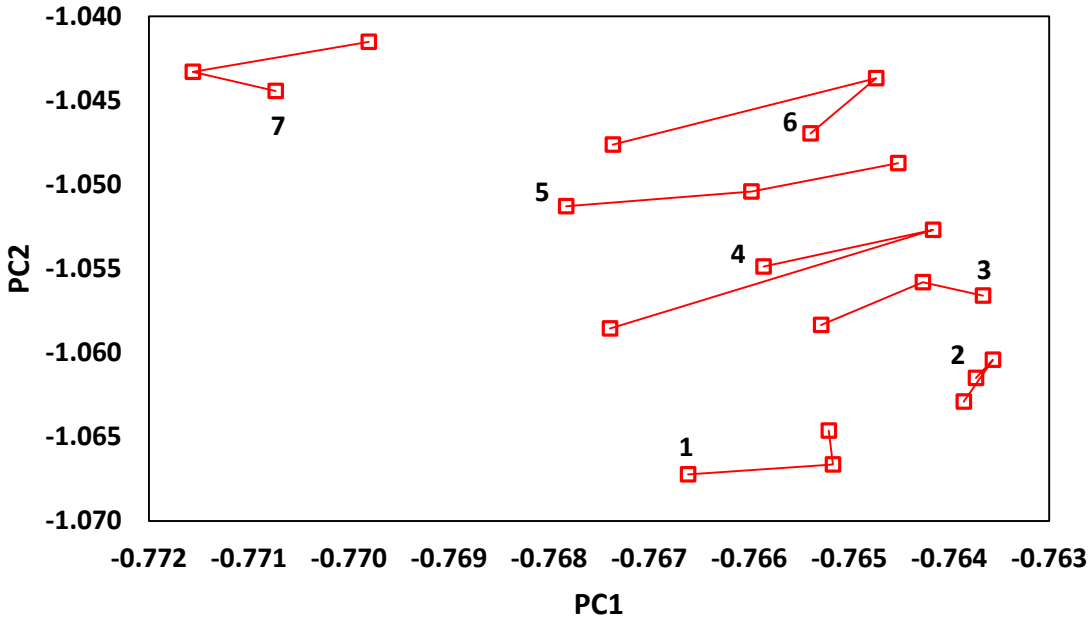


Figure 34. Expanded view of scores plot from scans of a single powdered 10% canola oil : white rice flour sample acquired with rotation. Groups of three scans acquired at 120° angles relative to each other are connected by lines. Numbers indicate the first scan of each group in temporal order.

5.1.4.2 Variability Among Replicate Scans of Multiple Samples

Results of replicate FT-NIR scans of ten samples each of 10% (w:w) canola oil: white rice flour assayed after tableting at 115 psi or 415 psi are shown in Figure 35. Both panels of Figure 35 are on the same scale and show an identical area of principal component space. Although the effect of sample rotation in reducing variation was less profound than in the case of a single sample, variation among scans of multiple samples was reduced by sample rotation on both principal component axes. Comparison of these scores plots also indicates rotation improved resolution among samples subjected to different forming pressures as evidenced by reduced overlap between different clusters of rotating samples (Figure 35b) than those of stationary ones (Figure 35a).

Standard deviation data for scores from sets of three scans of each of ten samples are shown in Table 16 and Figure 36. Sample rotation reduced the standard deviation of scores by 30% along Principal Component 1 and by 33% along Principal Component 2 for the set of low pressure (115 psi) sample tablets. Rotation also reduced the standard deviation of scores by 59% along Principal Component 1 and by 56% along Principal Component 2 for the set of high pressure (415 psi) sample tablets. The difference in the effect of rotation between the two forming pressures was likely the result of increased sample inhomogeneity induced by the higher pressure. Although scores from stationary scans of the high pressure samples were more variable, sample rotation eliminated this effect. Given the advantages of pressure for sample handling, the 415 psi forming pressure was selected for the quantitative ASL study.

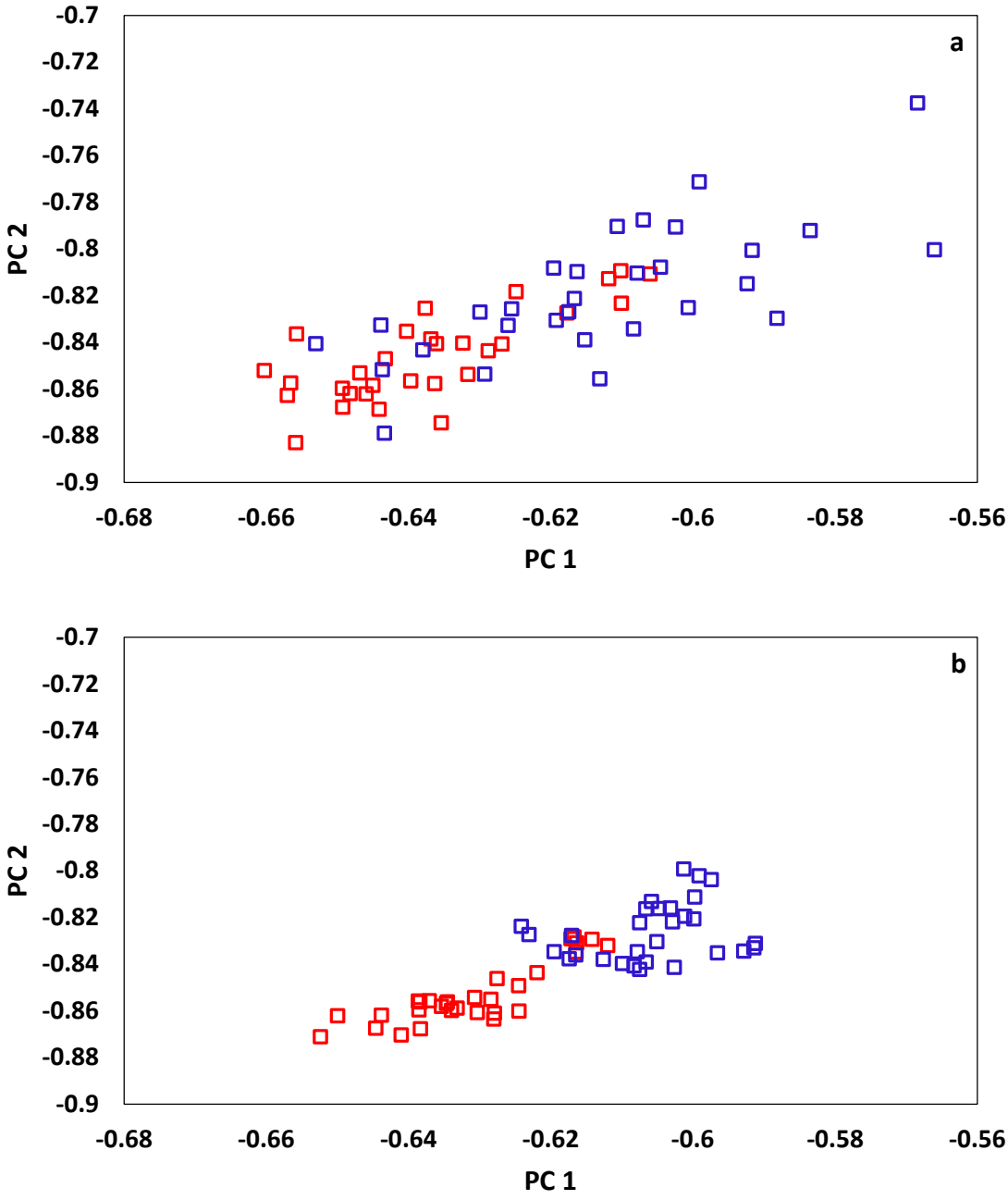


Figure 35. Scores plots of 10% canola oil: white rice flour samples: (a) Stationary; (b) Rotating [□ 115 psi Tablets; □ 415 psi Tablets].

Table 16. Standard deviation of scores of ten 10% canola oil : white rice flour samples.

3 Scans Each of 10 Samples	115 psi Tablets (n=30)		415 psi Tablets (n=30)	
	Stationary	Rotating	Stationary	Rotating
Principal Component 1	1.51×10^{-2}	1.06×10^{-2}	2.14×10^{-2}	8.72×10^{-3}
Principal Component 2	1.94×10^{-2}	1.29×10^{-2}	2.82×10^{-2}	1.23×10^{-2}

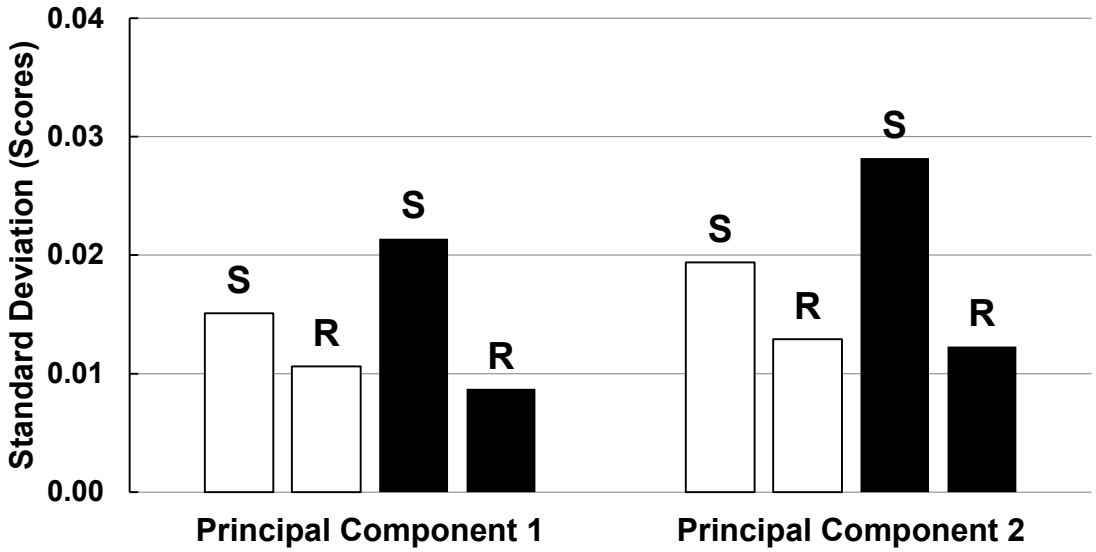


Figure 36. Standard deviation among scores of triplicate scans of ten 10% canola oil : white rice flour sample tablets [☐ 115 psi; ■ 415 psi; Stationary (S); Rotating (R)].

An expanded view of the rotating 415 psi sample scores is shown in Figure 37. Unlike scores from repeated scanning a single rotating tablet as shown in Figure 33 or from scanning multiple stationary tablets as shown in Figure 26, no trend was evident in the scores of multiple rotating tablets. The trending in earlier experiments may have been a temperature or moisture effect. Repeated scanning of a single sample increased the duration that the sample was atop the spectrometer, allowing greater shifts in temperature or moisture content of the sample between initial and final scans. Although scanning only three replicates per sample as in the qualitative canola oxidation study provided much less time for changes associated with these effects, repeated measurement of a stationary point was not complicated by shifts in scores due to

changes in the sample composition which arise from rotation of the sample during scanning. Thus, scores of replicate scans of each sample in Figure 24 and Figure 26 showed trending, while the initial placement for any particular sample in those scores plots depended upon its composition. The trending effect was not evident in scores of the multiple rotating sample tablets in Figure 37, likely due to the brief time each sample spent on the spectrometer and the fact that sample rotation gave rise to shifting composition effects which obscured trending.

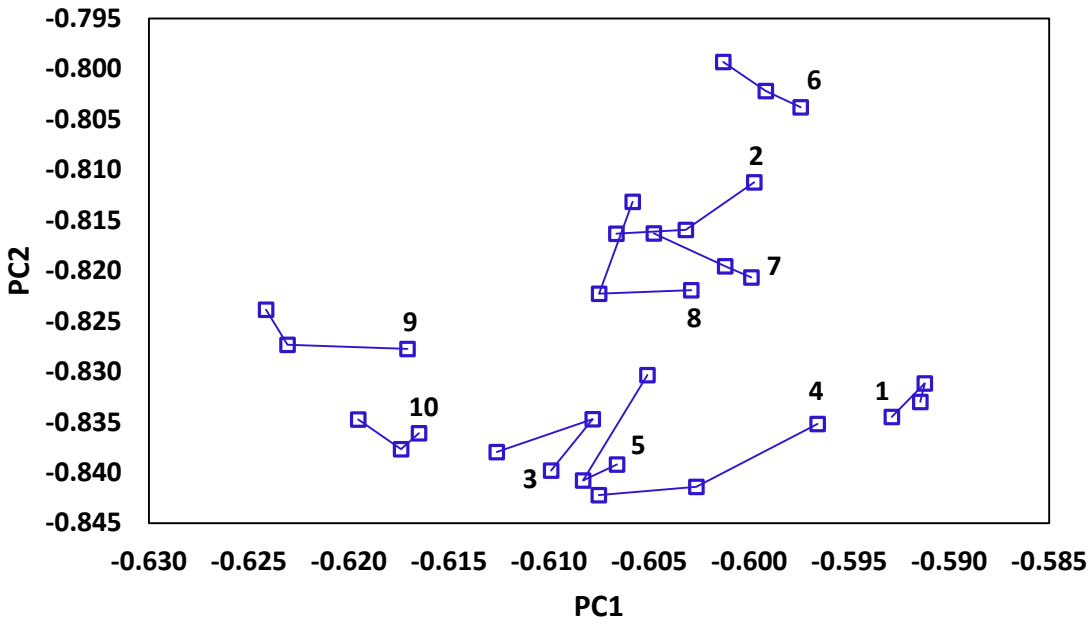


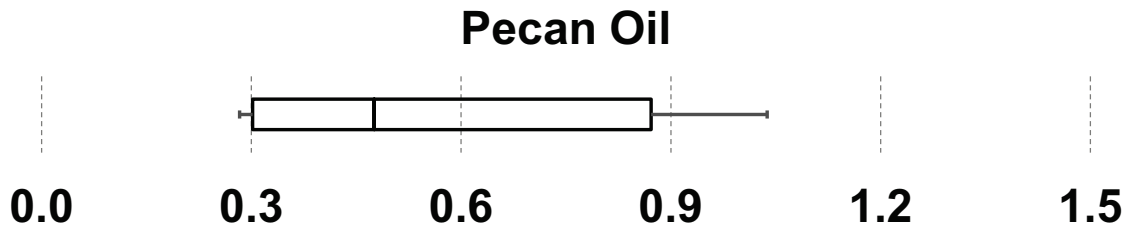
Figure 37. Expanded view of scores plot of rotating sample tablets pressed at 415 psi. Replicate scans made on each sample tablet are connected by lines. Numbers indicate the first scan of each of the ten sample tablets assayed in temporal order.

5.2 Quantitative Analysis of Lipid Oxidation: Accelerated Shelf Life Study

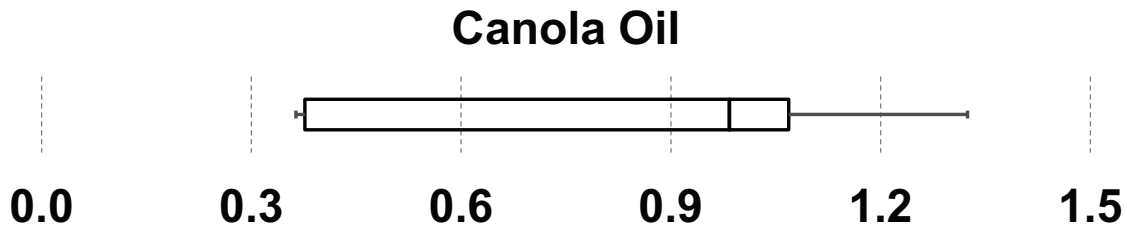
5.2.1 Experimental Results

5.2.1.1 Conjugated Dienes Reference Assay

The distribution of oxidation index values among pecan and canola oils indicated a disparity in extent of oxidation among oils despite similar incubation times (Figure 38). Pecan oil was more stable to oxidation. Initial (T0) canola oil samples had larger oxidation index values and the range of values during the course of the canola ASL was nearly 30% greater. Overall ranges for both oils skewed toward low levels of oxidation, likely due to the fact that neither was stripped of antioxidants and that incubations occurred at 40°C rather than higher temperatures.



Minimum	1st Quartile	2nd Quartile	3rd Quartile	Maximum
0.2831	0.3016	0.4756	0.8717	1.038



Minimum	1st Quartile	2nd Quartile	3rd Quartile	Maximum
0.3639	0.3764	0.9836	1.069	1.324

Figure 38. Box and whiskers plots showing distribution of oxidation index values for oils during the 40°C ASL study. Whiskers indicate first and fourth quartiles of data and boxes indicate second and third quartiles of data.

Plots of oxidation index versus incubation time for each oil are shown in Figure 39, which displays the average of the three samples taken at each time point with error bars indicating 95% confidence intervals.

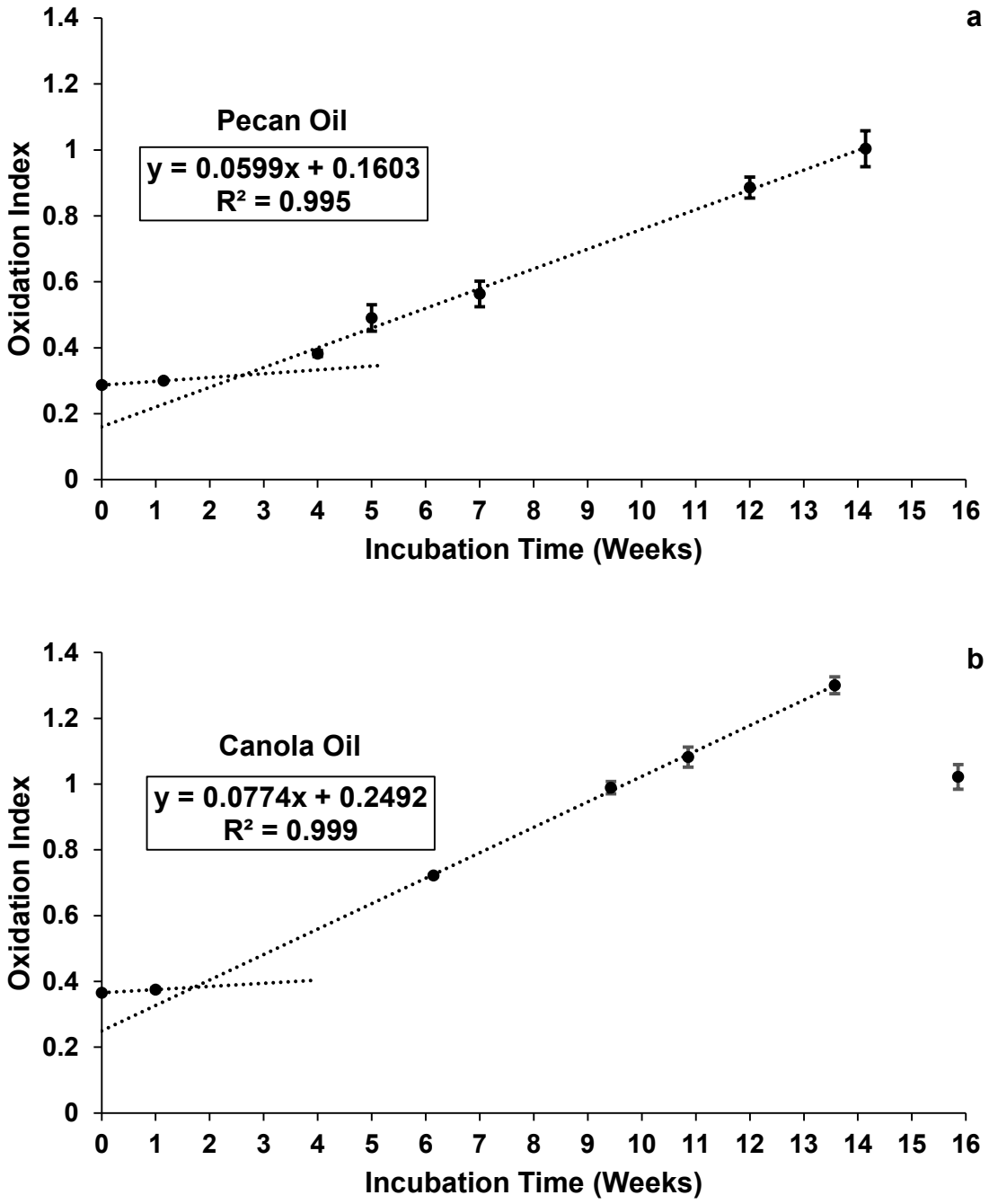


Figure 39. Oxidation index (conjugated dienes) for (a) pecan and (b) canola oils incubated at 40°C. Each point represents the average of three replicates drawn at that incubation time. Where error bars do not appear they are below the size of the marker.

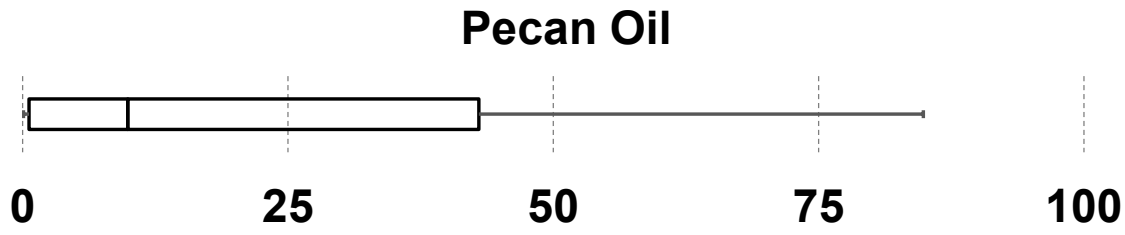
Klein's oxidation index has been reported to have a linear relationship with the duration (Babincova et al., 1999) and intensity (Mandal et al., 1978) of oxidizing conditions. Three points are clearly evident from the figure. First, the first two samples of each oil were taken during a lag period prior to the propagation phase accompanied by development of conjugated dienes at a constant rate. The lag period for pecan oil was longer than that for canola oil, a consequence of different fatty acid composition and different levels of antioxidants in these oils. Second, the rate at which conjugated dienes were generated during the propagation phase was nearly 30% higher in canola oil based on the ratio of slopes of the propagation phase lines. Third, while no clear maximum was reached for pecan oil, oxidation index values for canola oil peaked during the study as samples at the final time point taken sixteen days after the previous assay were far below the growth line. This decrease was consistent with the conclusion that conversion of conjugated dienes to secondary oxidation products was reached in canola.

5.2.1.2 Lipid Hydroperoxides Reference Assay

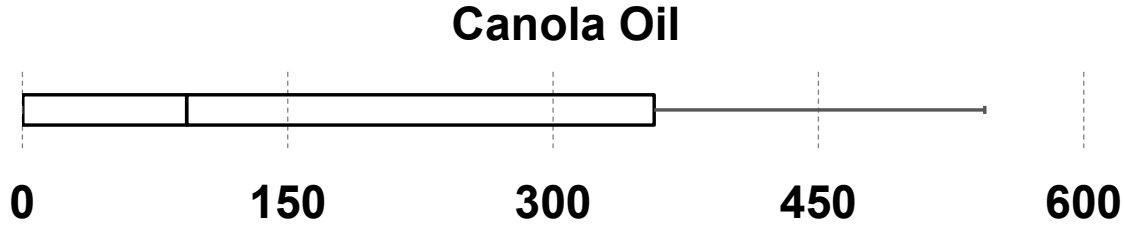
The distribution of peroxide values among pecan and canola oils during the ASL, shown in Figure 40, indicated a large disparity between the two oils. The range of peroxide values observed in canola oil was nearly 6.5 times that of pecan oil. As was observed for conjugated dienes, peroxide values in both sets of oil samples were skewed towards lower levels of oxidation under the conditions investigated.

Plots of peroxide values versus incubation time for each oil are shown in Figure 41, which displays the average of the three samples taken at each time point with error bars indicating 95% confidence intervals. Peroxide values for pecan oil (Figure 41a) were less than one quarter of the overall range for the first five of the seven time points evaluated. Variability increased as peroxide values rose in pecan oil. The latter observation was even more pronounced in canola oil (Figure 41b) where data from the last two time points were highly variable. Unlike the case in pecan oil, in canola oil the changes between peroxide values of the final two time points appeared

to level off and their confidence interval ranges largely overlapped. This reinforces the inference from the oxidation index data that canola oil but not pecan oil reached the point at which levels of primary lipid oxidation products began to level off or drop, likely as secondary products arose.



Minimum	1st Quartile	2nd Quartile	3rd Quartile	Maximum
0.07592	0.6031	9.909	43.00	84.69



Minimum	1st Quartile	2nd Quartile	3rd Quartile	Maximum
0.04033	0.2195	92.85	357.2	544.0

Figure 40. Box and whiskers plots showing distribution of peroxide values (meq/Kg) for oils during the 40°C ASL study. Whiskers indicate first and fourth quartiles of data and boxes indicate second and third quartiles of data.

5.2.1.3 Non-Volatile Carbonyl Products Reference Assay

Analysis of carbonyl secondary products of lipid oxidation by the modified DNPH assay indicated that incubations at 40°C for just over fourteen weeks in pecan oil or over fifteen weeks in canola oil were insufficient to generate substantial amounts of these products. No carbonyls were observed in chromatograms of any of the pecan oil samples, while among canola oil samples short chain saturated carbonyls developed only in the final two time points of the study (Figure 42).

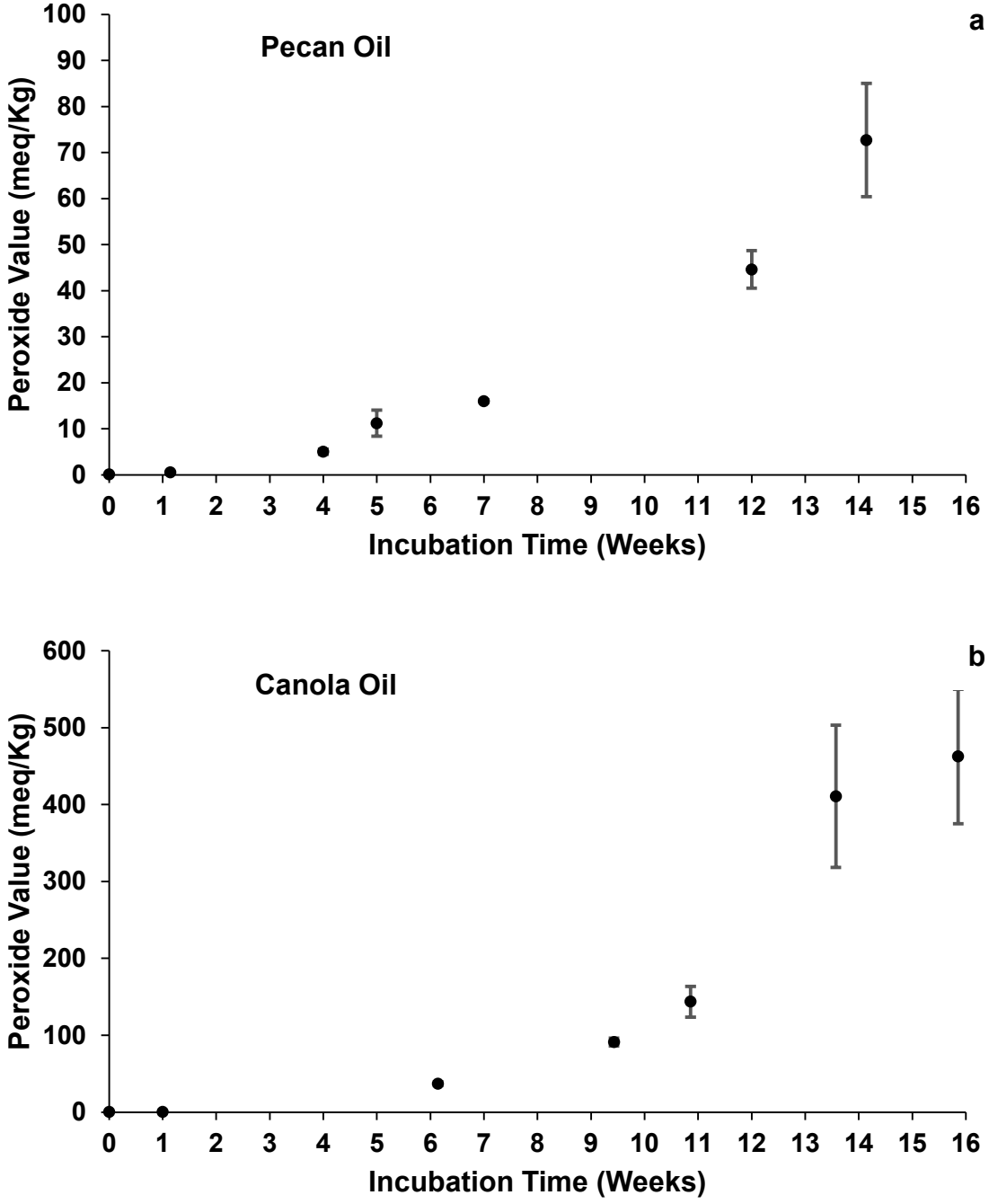


Figure 41. Peroxide values for (a) pecan and (b) canola oils incubated at 40°C. Each point represents the average of three replicates drawn at that incubation time. Where error bars do not appear they are below the size of the marker.

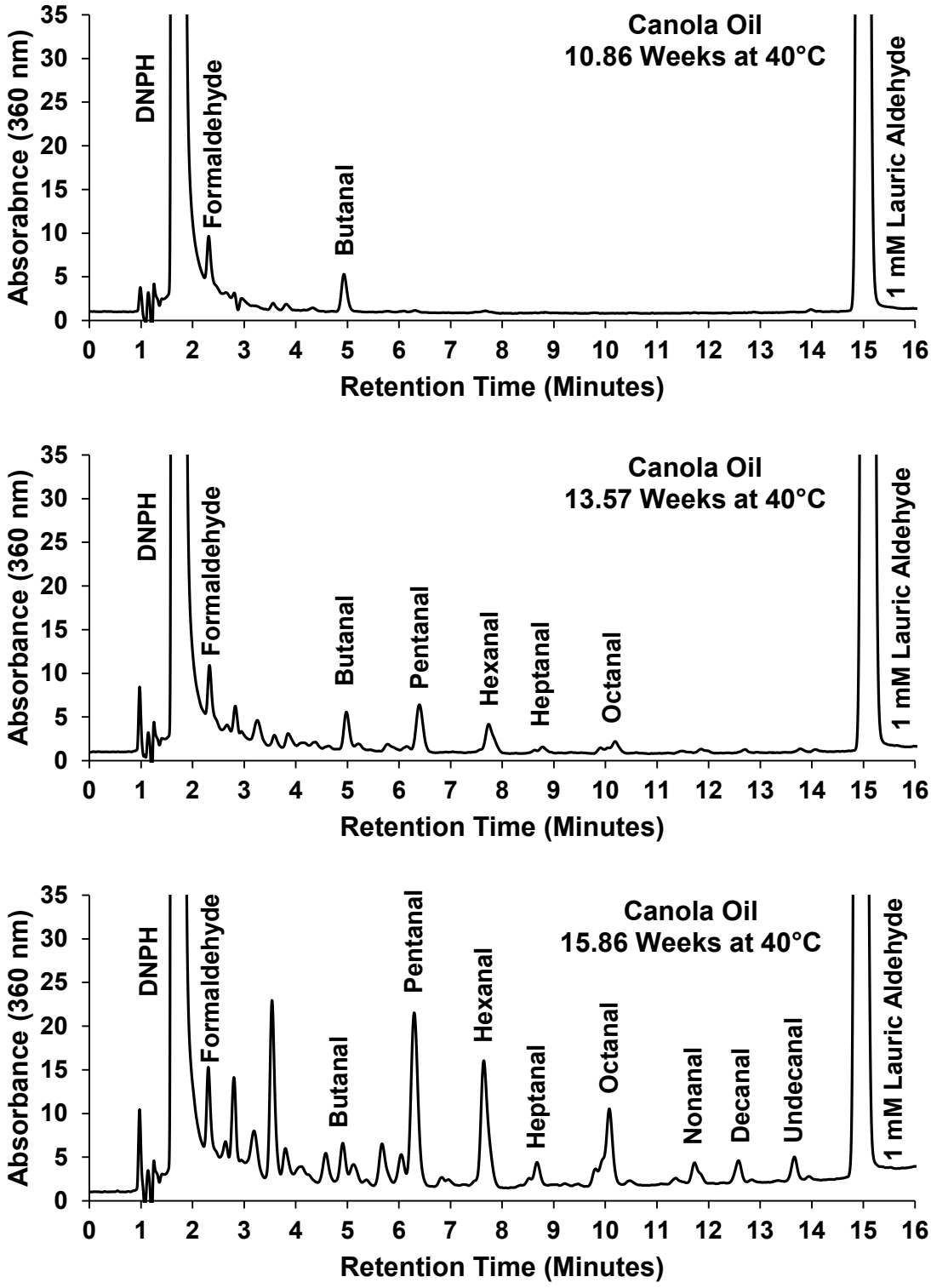


Figure 42. 360 nm chromatograms following DNPH derivatization of canola oil incubated at 40°C for 10.86 weeks (top), 13.57 weeks (center) and 15.86 weeks (bottom).

The conditions for the DNPH assay relied upon lauric aldehyde as a valid internal standard for saturated aldehydes, while less reactive unsaturated aldehydes were monitored using 233 nm for 2-enals and 270 nm for 2, 4-dienals. Incorporation of the internal standard indicated the assay worked properly for saturated aldehydes, as peaks for the hydrazone derivative of lauric aldehyde were consistently observed throughout all blanks and samples. Saturated aldehydes generated from sample incubation were only observed in the final two time points of the canola oil incubation.

Although a small peak corresponding to the retention time for butanal was observed at almost 11 weeks of incubation, it did not arise from oxidation as the chromatogram shown in Figure 42 (top) was identical to the blank. Interestingly, butanal was not found to increase in later time points, but a many products arose at shorter retention times in addition to successively longer chain saturated aldehydes (Figure 42 center, bottom). Casale et al. (2007) investigated the kinetics of aldol condensation in C2-C8 aliphatic aldehydes and determined butanal to be the most reactive. Our observations could indicate butanal was not a significant product of oxidation, evaporated during sample incubation, or underwent side reactions during the DNPH assay.

While many of the products eluting prior to butanal were also observed in the respective blanks at diminished concentrations, sample chromatograms were much noisier and contained a peak at 2.8 minutes (Figure 42 bottom) not found in the blank. Yao (2015) also found large amounts of early eluting products during studies of lipid oxidation using a similar DNPH assay. These could be natural products of lipid oxidation or side products generated during the DNPH assay. None of the seven longer chain saturated aldehydes observed in the final week samples of canola (Figure 42 bottom) were observed in the blank.

Monitoring at 233 nm indicated characteristic peaks for each oil between 16 and 26 minutes (Figure 43). Peak areas among sample batches correlated with conjugated diene levels.

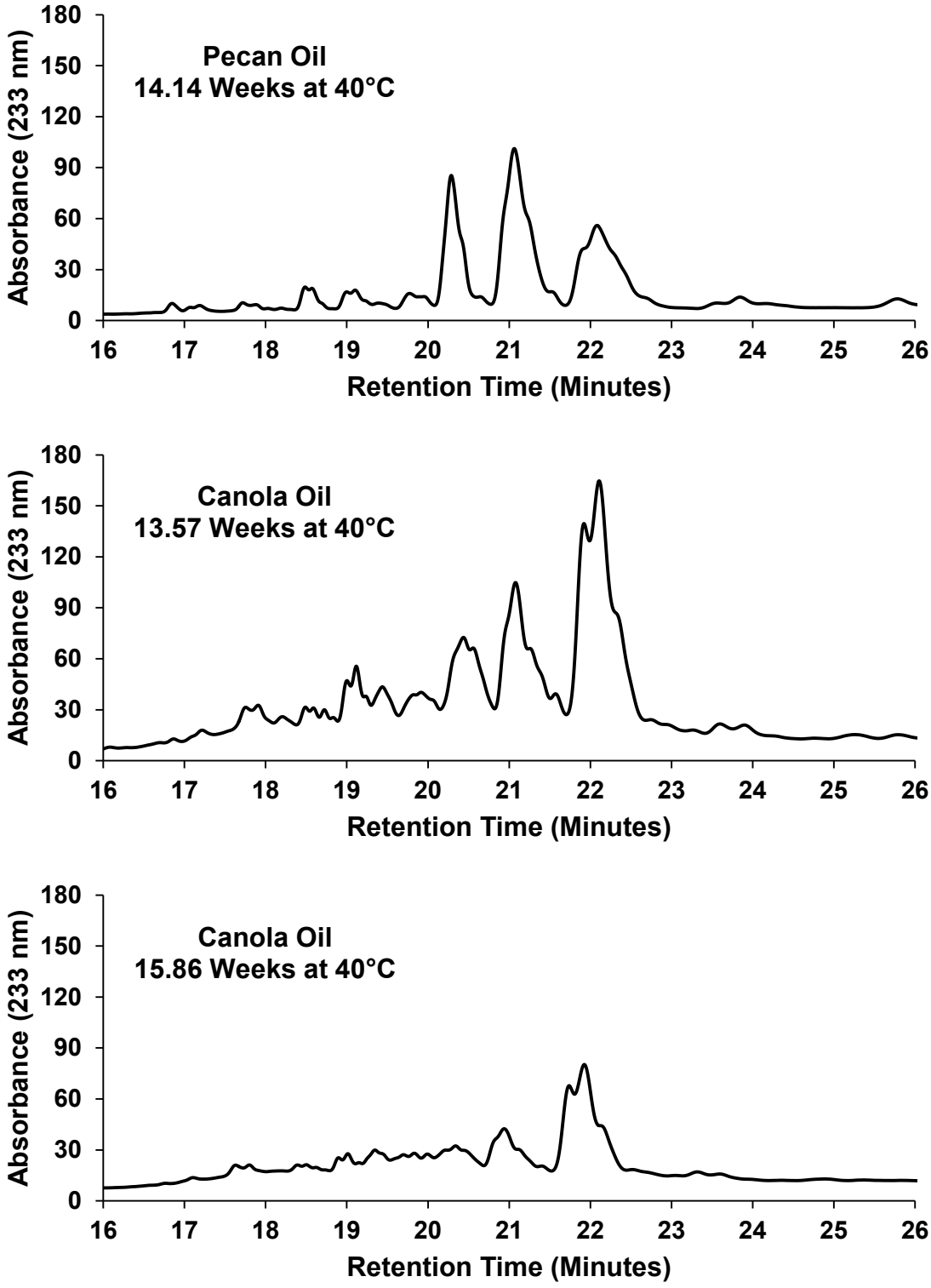


Figure 43. 233 nm chromatograms of oils incubated at 40°C following DNPH derivatization: pecan oil at 14.14 weeks (top); canola oil at 13.57 weeks (center); and canola oil at 15.86 weeks (bottom).

Given the retention times of these products, they were tentatively identified as core conjugated dienes attached to triacylglycerols. Chromatograms of samples containing the maximal amount of conjugated dienes observed for pecan oil (Figure 43 top) and canola oil (Figure 43 center) indicated a greater amount of a core product eluting between 22 and 23 minutes in canola. This discrepancy among pecan and canola oils was consistent throughout the ASL, thus it was likely attributable to differences among the respective fatty acid compositions of the two oils. Typical compositions for the two are shown in Table 17.

Table 17. Typical fatty acid composition of pecan and canola oils.

Fatty Acid	Pecan Oil*	Canola Oil**	Pecan Oil***	Canola Oil***
Palmitic (16:0)	5.2%	5.2%		
Stearic (18:0)	2.7%	4.4%		
Saturated	7.9%	9.6%	9.5%	9.6%
Oleic (18:1)	64.6%	59.5%		
Monounsaturated	64.6%	59.5%	52.0%	59.7%
Linoleic (18:2)	24.4%	18.8%		
Linolenic (18:3)	2.2%	11.9%		
Polyunsaturated	26.6%	30.7%	38.5%	30.7%
(Other)	0.9%	0.2%	-	-

* (Toro-Vazquez et al., 1999) Average of values from three growing regions.
** (Kostik et al., 2013) Values for high oleic variety.
*** ("Benefits of Cooking With Pecan Oil", 2019)

Toro-Vazquez et al. (1999) observed that the content of oleic acid in pecan oil dropped while those of linoleic and linolenic acids rose with tree age, with linoleic acid content significantly greater than linolenic acid. Comparison of values from the first two columns of Table 17 indicates the largest differences in fatty acid composition between pecan oil and high oleic canola oil are attributable to these three unsaturated fatty acids. Values in the final two columns of Table 17 are from the website of Kinloch Plantation Products, LLC, the manufacturer of the pecan oil used in this study. The canola oil values are identical to those of Kostik et al. (2013), while those for pecan oil differ greatly from Toro-Vazquez et al. (1999) in the amount of unsaturated fatty acids. This indicates that the pecan oil used herein had lower contents of oleic

and linolenic acids but a much higher content of linoleic acid than the high oleic canola oil assayed herein. Differences in conjugated dienes evident from chromatograms of the two oils at 233 nm thus reflect differences in fatty acid composition. The difference in oxidation rates of the two oils was attributable to the very high level of natural antioxidants present in pecans, 179.4 μmol Trolox equivalents (TE)/g (Wu et al., 2004) relative to those typically found in plant oils [i.e. 2.20 μmol TE/g in soybean oil, 1.79 μmol TE/g in extra virgin olive oil, and 1.17 μmol TE/g in sunflower oil (Pellegrini et al., 2003)].

The changes in chromatograms of canola oil from the penultimate (Figure 43 center) and final (Figure 43 bottom) time points of the ASL also reflected the observations from the conjugated dienes assay (Figure 39b). Reduction of conjugated dienes in the final canola oil samples corresponded with the increase in secondary carbonyl products of lipid oxidation (Figure 42 bottom).

Results may have been attributable to gentle incubation conditions, to evaporation of volatile aldehydes during incubation, and/or to the presence of endogenous or added antioxidants in the commercial oils. The incubation periods used for these oils at 40°C were insufficient to reach the later stages of lipid oxidation in which large amounts of carbonyl secondary products arose. These observations were also interesting in light of alternate competing pathways of lipid oxidation being studied in this laboratory. Carbonyl products typically result from scission reactions of lipid alkoxyl radicals. A strong proton donating source is necessary to stabilize scission products and drive the reaction forward (Schaich, 2005). In oxidizing oils with randomly oriented triacylglycerol molecules, abstractable hydrogens are not readily available. Peroxyl radicals add preferentially to lipid double bonds, generating epoxides instead of hydroperoxides. Preliminary testing of oils from this experiment indicated the presence of elevated levels of epoxides. These shifting reaction pathways could explain why carbonyls were not detected under the conditions investigated. More detailed analyses will be necessary in future studies. Without available data from the carbonyl reference assay, NIR analysis was not performed for carbonyls.

5.2.1.4 Near Infrared Spectroscopy

All 945 original reflectance spectra of samples made with each oil are shown in Figure 44. The broad bands in each figure are the result of baseline shifting of individual spectra due to variation in the intensity of light returning to the detector among different samples. The most intense spectra in Figure 44a are triplicate scans accidentally acquired on a sample that was not placed on a borosilicate glass disc. This shows the damping effect on spectral intensity that substrates such as the borosilicate glass discs used here can have in transflectance spectroscopy. Replicate spectra on this sample were immediately reacquired with the sample on a borosilicate glass disc and that data was used for chemometric modeling.

Figure 44 shows all of the individual scans made during the ASL. To create chemometric models relating NIR and chemical assay data, the average spectrum of the forty five scans acquired for each sample batch were used, following the application of various pretreatments, to compare with the values of conjugated dienes and peroxides obtained for that batch from the applicable reference assay. Examples of spectra for the twenty one batches of pecan oil samples and canola oil samples following normalization (SNV) and smoothing (9 point Savitzy-Golay gap 2) pretreatment are shown in Figure 45. The models developed herein relied upon correlation of the minor changes in these spectra with chemical assays of the corresponding samples.

5.2.2 Full Spectrum Predictive NIR Models for Lipid Oxidation Parameters

Chemometric models using partial least squares regression (PLSR) were developed for conjugated diene content [expressed as the oxidation index (Klein, 1970)] and peroxide values (meq/kg) in pecan and canola oils subjected to accelerated shelf life incubation and coated on white rice flour at 7.5% oil by weight. Models were also assessed using PCR; however, the fitness of those was comparable to or worse than PLSR models. Principal components analysis is useful in qualitative analyses for sample identification (such as in the lipid discrimination, forming pressure discrimination, rotation or qualitative ASL studies herein) or where quantitative

reference assay data is lacking. PLS factors, which are derived using the quantitative reference assay data, are more easily interpreted than principal components in quantitative analyses. Given this increased difficulty in interpretation and the lack of any improvement in PCR models over PLSR models, analysis was constrained to PLSR.

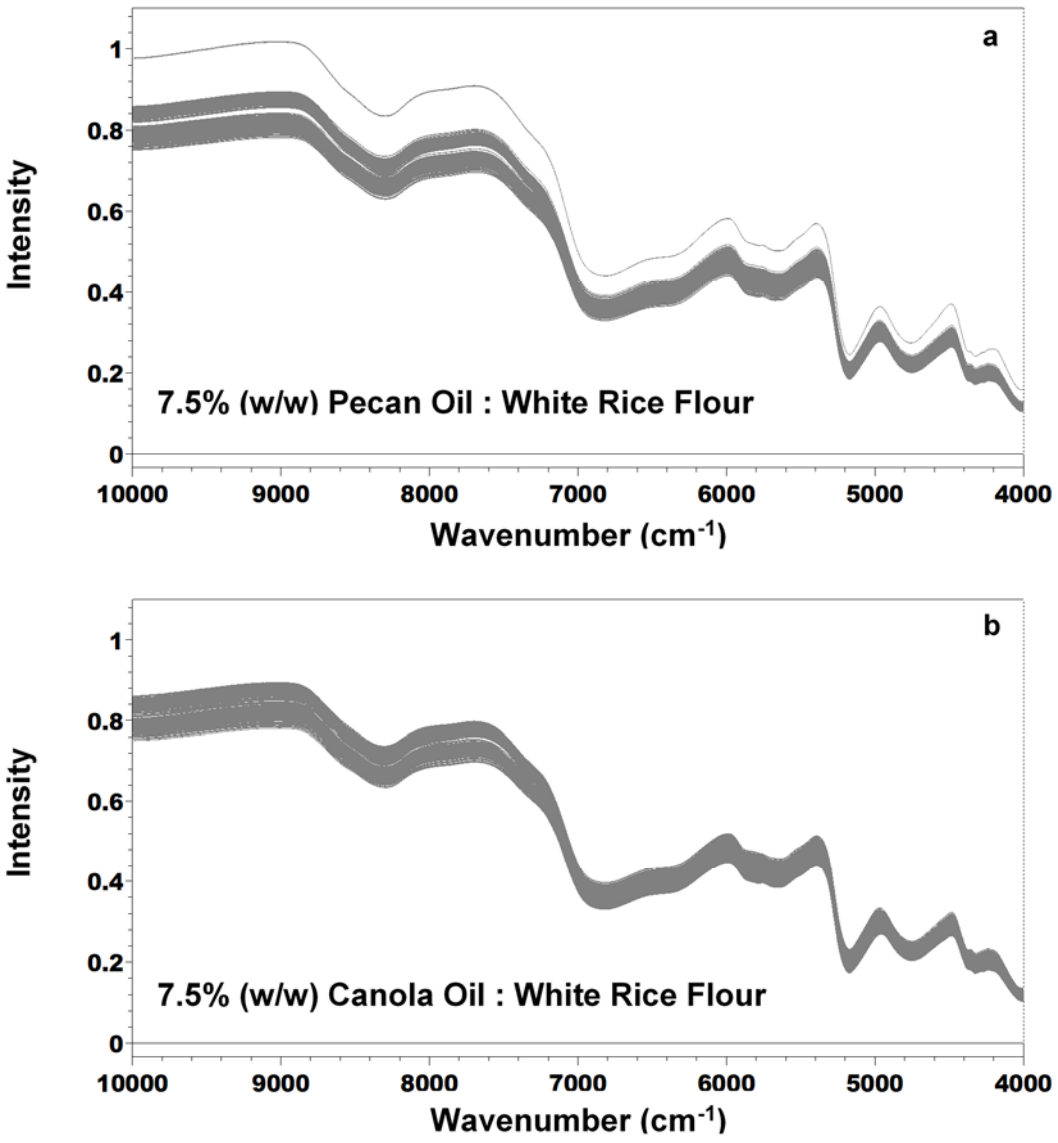


Figure 44. Original NIR reflectance spectra of 7.5% (w/w) nominal mixtures of oils from the ASL study with white rice flour: (a) Pecan oil samples and (b) Canola oil samples.

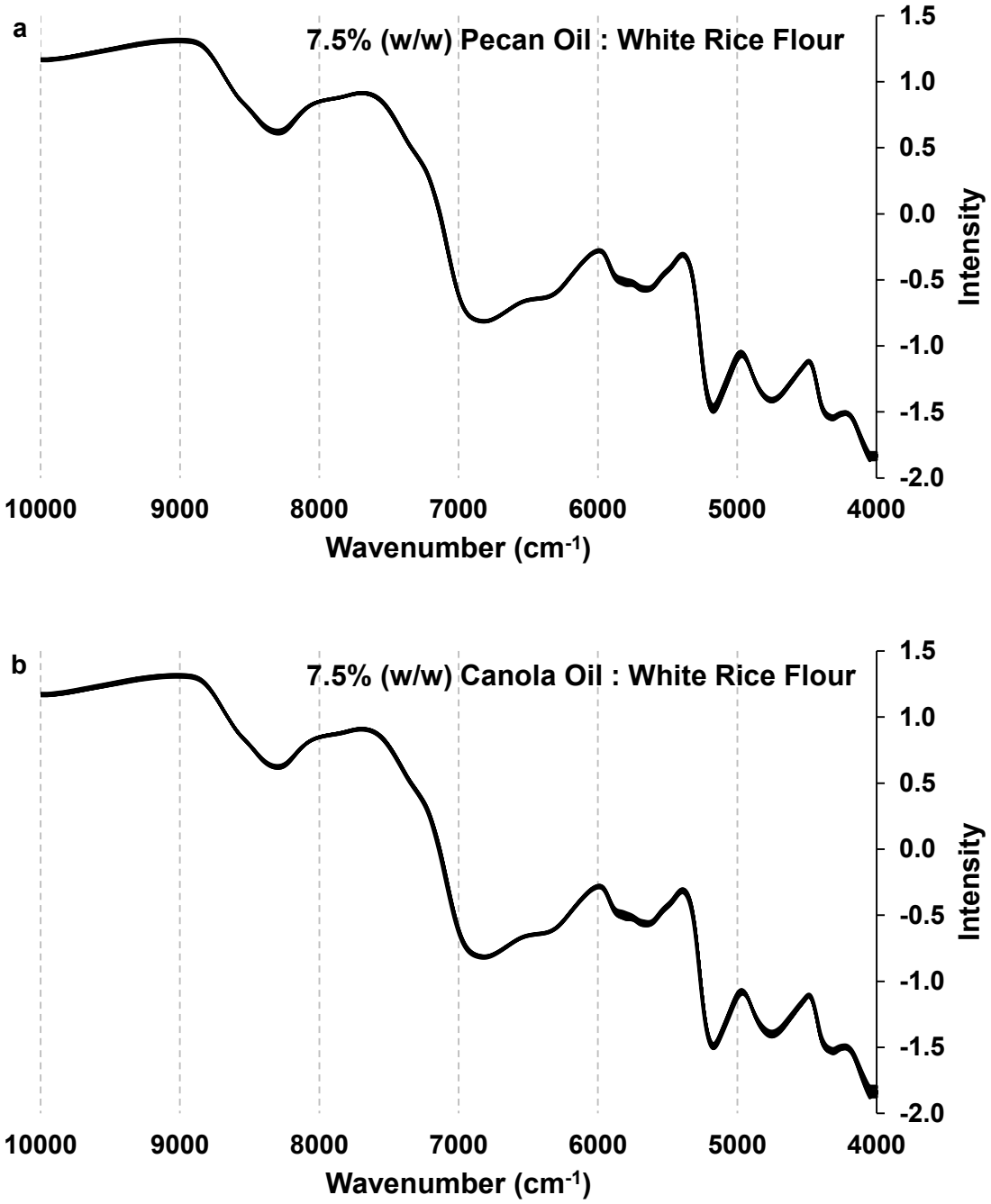


Figure 45. Normalized and smoothed average NIR spectra for each of the twenty one 7.5% oil : white rice flour batches from ASL studies of (a) Pecan oil and (b) Canola oil.

PLSR models for conjugated dienes and lipid hydroperoxides were developed for samples made with pecan oil and canola oil as well as for a data set combining both sample types. Models relied upon reference assay data for the twenty-one samples of each oil from the ASL study as well as the average NIR spectrum of the 7.5% oil : white rice flour mixture made with each oil sample. Each average spectrum was a composite of the forty-five spectra (triplicate scans of fifteen pressed tablets acquired with rotation) acquired for that oil sample. Initial models were generated by applying PLSR to the full NIR spectrum (4000 - 10,000 cm^{-1}).

5.2.2.1 NIR Assays of Conjugated Dienes (Oxidation Index)

NIR models for conjugated dienes (expressed as the oxidation index (Klein, 1970)) using 7.5% (w/w) pecan oil mixed with white rice flour were generally very poor under the conditions studied, while those using similar samples incorporating 7.5% (w/w) canola oil were slightly better. A data set combining both sets of samples provided models with quality nearly as good as those for canola oil samples. A number of factors were identified to explain these results and indicate possible adjustments in sample handling and analysis that can improve the technique.

5.2.2.1.1 Pecan Oil Samples

Q-values for models constructed using various pretreatments on the spectral data of samples made with pecan oil are shown in Table 18. Outlier analysis of these full spectra models was complicated by a fair amount of scatter in residual values. This was likely attributable to the relatively low level of oxidation in many of the samples given the stability of pecan oil (see Figure 38). Candidates for exclusion included the T0 samples as well as those from the final week (Week 14) given that these samples capped the data and thus had high leverage in addition to scatter among residuals.

As seen in Table 18, model quality improved by removing the T0 samples but dropped when the final week samples were excluded, likely because of low levels of oxidation among the

remaining samples. Eliminating the T0 samples reduced scattering at the low end of the data, while eliminating samples from the final week discarded the most oxidized samples.

Table 18. Q-values for models of conjugated dienes in 7.5% pecan oil samples.

Q-Value	PLS Factors	Normalization (SNV)	Smoothing (9 Point S-G Gap 2)	Derivatization	Samples Excluded from Model	n
0.4445	3	-	-	-	-	21
0.5096	2	+	-	-	-	21
0.5097	2	+	+	-	-	21
0.5288	2	+	+	1st	-	21
0.4705	2	+	+	2nd	-	21
0.5807	5	-	-	-	T0	18
0.6182	2	+	-	-	T0	18
0.6182	2	+	+	-	T0	18
0.6266	2	+	+	1st	T0	18
0.6183	2	+	+	2nd	T0	18
0.3945	2	-	-	-	Week 14	18
0.4494	2	+	-	-	Week 14	18
0.4489	2	+	+	-	Week 14	18
0.4336	2	+	+	1st	Week 14	18
0.4282	2	+	+	2nd	Week 14	18
0.5153	5	-	-	-	T0 and Week 14	15
0.5501	2	+	-	-	T0 and Week 14	15
0.5502	2	+	+	-	T0 and Week 14	15
0.5465	2	+	+	1st	T0 and Week 14	15
0.5433	2	+	+	2nd	T0 and Week 14	15

The best full spectrum model was obtained by discarding only the T0 samples and applying normalization, smoothing and first derivative transformation of the spectral data. This full spectrum model afforded a Q-value of 0.6266.

5.2.2.1.2 Canola Oil Samples

Q-values for models constructed using various pretreatments on the spectral data of samples made with canola oil are shown in Table 19. Although canola oxidized to a greater extent than pecan, outlier analysis remained difficult due to scatter among the data. Again, initial and final time points were candidates for exclusion given their increased leverage in addition to scatter among their residuals.

Table 19. Q-values for models of conjugated dienes in 7.5% canola oil samples.

Q-Value	PLS Factors	Normalization (SNV)	Smoothing (9 Point S-G Gap 2)	Derivatization	Samples Excluded from Model	n
0.5813	4	-	-	-	-	21
0.5371	4	+	-	-	-	21
0.5364	4	+	+	-	-	21
0.5995	4	+	+	1st	-	21
0.5839	4	+	+	2nd	-	21
0.4861	3	-	-	-	T0	18
0.5454	3	+	-	-	T0	18
0.5455	3	+	+	-	T0	18
0.5851	3	+	+	1st	T0	18
0.5524	3	+	+	2nd	T0	18
0.6349	4	-	-	-	Week 15	18
0.6018	4	+	-	-	Week 15	18
0.6332	5	+	+	-	Week 15	18
0.6712	3	+	+	1st	Week 15	18
0.6976	3	+	+	2nd	Week 15	18
0.6438	5	-	-	-	T0 and Week 15	15
0.6778	4	+	-	-	T0 and Week 15	15
0.7252	6	+	+	-	T0 and Week 15	15
0.6949	3	+	+	1st	T0 and Week 15	15
0.6801	3	+	+	2nd	T0 and Week 15	15

While exclusion of the T0 samples alone failed to improve model quality, removal of only the final week samples did, while the greatest improvement resulted from removing both groups of samples. This result was interesting as it is typical for model quality to degrade when using fewer samples. The best such full spectrum model was obtained by applying normalization and smoothing pretreatments and had a Q-value of 0.7252.

5.2.2.1.3 Combined Samples

PLS models of conjugated dienes were also obtained using a combined data set of pecan and canola samples over the full spectral range. A summary of the best models for each sample set of individual oils is shown in Table 20. These results indicated that normalization, smoothing and first derivative pretreatment resulted in the best model obtained for pecan oil samples in addition to accounting for half of the top models for both the pecan and canola oil samples. Normalization and 9 point Savitzky-Golay (Gap 2) smoothing provided the best model for canola oil samples and also the second best model for pecan oil samples. Accordingly, each pretreatment was evaluated for all permutations of T0 and final week sample exclusions of both oils in the combined data set.

Table 20. Best individual oil conjugated diene models for each sample set.

Q-Value	PLS Factors	Normalization (SNV)	Smoothing (9 Point S-G Gap 2)	Derivatization	Samples Excluded from Model	n	
0.6266	2	+	+	1st	T0	18	Pecan
0.5502	2	+	+	-	T0 and Week 14	15	
0.5288	2	+	+	1st	-	21	
0.4494	2	+	-	-	Week 14	18	
0.7252	6	+	+	-	T0 and Week 15	15	Canola
0.6976	3	+	+	2nd	Week 15	18	
0.5995	4	+	+	1st	-	21	
0.5851	3	+	+	1st	T0	18	

Results of models prepared using normalization with smoothing and first derivative pretreatment are shown in Table 21 ranked by Q-value. Each of the best five models eliminated the final week samples for canola and all but one also excluded the final week samples for pecan. Most excluded the T0 samples for canola as well. Overall, improved quality was observed in conjugated diene models using this pretreatment as samples were excluded.

Table 21. Combined data models of conjugated dienes using SNV and first derivative pretreatment ranked by Q-value.

Q-Value	PLS Factors	Pecan Samples Excluded		Canola Samples Excluded		n
		T0	Week 14	T0	Week 15	
0.6946	3	X	X	X	X	30
0.6795	6		X		X	36
0.6774	5		X	X	X	33
0.6665	3			X	X	36
0.6393	6	X	X		X	33
0.6282	3	X		X		36
0.6202	3	X	X	X		33
0.6139	3			X		39
0.6131	5				X	39
0.6102	3		X	X		36
0.5790	5	X			X	36
0.5529	3		X			39
0.5472	3	X		X	X	33
0.5440	5					42
0.5355	6	X	X			36
0.5145	6	X				39

Results of models prepared using normalization and smoothing pretreatment are shown in Table 22 ranked by Q-value. Although the change in pretreatment shifted the order of sample sets relative to model quality, the general trend towards improvement with exclusion of samples, particularly with respect to initial and final canola samples, was preserved. This pretreatment resulted in the best overall model for the combined data set by excluding the final weeks of each oil as well as the initial canola oil samples, with a Q-value of 0.7214.

Table 22. Combined data models of conjugated dienes using SNV and smoothing pretreatment ranked by Q-value.

Q-Value	PLS Factors	Pecan Samples Excluded		Canola Samples Excluded		n
		T0	Week 14	T0	Week 15	
0.7214	5		X	X	X	33
0.6966	5	X	X	X	X	30
0.6734	5			X	X	36
0.6616	5		X		X	36
0.6375	4		X	X		36
0.6202	4			X		39
0.6163	5	X	X		X	33
0.6159	3	X		X		36
0.6124	5				X	39
0.5940	4	X	X	X		33
0.5675	5	X		X	X	33
0.5659	5	X			X	36
0.5589	5		X			39
0.5380	5					42
0.5182	5	X	X			36
0.4909	5	X				39

5.2.2.2 NIR Assays of Lipid Hydroperoxides

Full spectrum NIR models of lipid hydroperoxides constructed using either 7.5% (w/w) pecan or canola oils mixed with white rice flour were found to be worse than those for conjugated dienes under the conditions studied. Unlike the case with conjugated dienes, use of a combined data set improved models relative to those of individual oils. Thus, sample size appears to have been more critical in the case of peroxides than conjugated dienes.

5.2.2.2.1 Pecan Oil Samples

Q-values for models constructed using various pretreatments on the spectral data of samples made with pecan oil are shown in Table 23. Similar to models for conjugated dienes, quality improved by removing the T0 samples but dropped when the final week samples were excluded. Again, the likely reason was low levels of oxidation among the set of pecan oil

samples. Eliminating the T0 samples reduced scattering at the low end of the data, while eliminating samples from the final week discarded the most oxidized samples.

Table 23. Q-values for models of peroxide values in 7.5% pecan oil samples.

Q-Value	PLS Factors	Normalization (SNV)	Smoothing (9 Point S-G Gap 2)	Derivatization	Samples Excluded from Model	n
0.3949	2	-	-	-	-	21
0.4575	2	+	-	-	-	21
0.4577	2	+	+	-	-	21
0.4648	2	+	+	1st	-	21
0.3041	2	+	+	2nd	-	21
0.4298	2	-	-	-	T0	18
0.5029	2	+	-	-	T0	18
0.5030	2	+	+	-	T0	18
0.5027	2	+	+	1st	T0	18
0.4702	2	+	+	2nd	T0	18
0.3533	2	-	-	-	Week 14	18
0.3934	2	+	-	-	Week 14	18
0.3936	2	+	+	-	Week 14	18
0.3548	2	+	+	1st	Week 14	18
0.3609	2	+	+	2nd	Week 14	18
0.4139	2	-	-	-	T0 and Week 14	15
0.4410	2	+	-	-	T0 and Week 14	15
0.4412	2	+	+	-	T0 and Week 14	15
0.4258	2	+	+	1st	T0 and Week 14	15
0.4293	2	+	+	2nd	T0 and Week 14	15

The best model, obtained by excluding only the T0 samples and pretreating spectra by normalization and smoothing, afforded a Q-value of only 0.5030. Comparison of these models with those for conjugated dienes (Table 18) using corresponding sample sets and pretreatments indicated a general reduction in quality among peroxide value models.

5.2.2.2.2 Canola Oil Samples

Q-values for models constructed using various pretreatments on the spectral data of samples made with canola oil are shown in Table 24. Again, comparison of these models with those for conjugated dienes (Table 19) indicated peroxide value models were lower in quality than corresponding models for conjugated dienes. However, unlike either peroxide models of pecan samples (Table 23) or conjugated diene models of canola samples (Table 19), removal of T0 or final week samples generated little if any improvement in model quality. Models were sufficiently poor that the highest Q-value (0.5264) was observed when using all samples with no pretreatments, a highly suspect result as in such a model baseline shifts between different samples were not corrected. The best quality model with pretreated spectra was obtained by dropping T0 samples and applying normalization and first derivative transformation (Q = 0.4692). This was slightly improved over a model which dropped the final week samples and used normalization with smoothing (Q = 0.4560).

Table 24. Q-values for models of peroxide values in 7.5% canola oil samples.

Q-Value	PLS Factors	Normalization (SNV)	Smoothing (9 Point S-G Gap 2)	Derivatization	Samples Excluded from Model	n
0.5264	4	-	-	-	-	21
0.4441	3	+	-	-	-	21
0.4434	3	+	+	-	-	21
0.4444	3	+	+	1st	-	21
0.4386	3	+	+	2nd	-	21
0.3748	4	-	-	-	T0	18
0.4458	2	+	-	-	T0	18
0.4449	2	+	+	-	T0	18
0.4692	2	+	+	1st	T0	18
0.4624	2	+	+	2nd	T0	18
0.4103	4	-	-	-	Week 15	18
0.4556	3	+	-	-	Week 15	18
0.4560	3	+	+	-	Week 15	18
0.4476	3	+	+	1st	Week 15	18
0.3850	3	+	+	2nd	Week 15	18
0.4064	2	-	-	-	T0 and Week 15	15
0.4308	2	+	-	-	T0 and Week 15	15
0.4302	2	+	+	-	T0 and Week 15	15
0.4140	2	+	+	1st	T0 and Week 15	15
0.4214	1	+	+	2nd	T0 and Week 15	15

5.2.2.2.3 Combined Samples

PLS models of peroxide values were also obtained using a combined data set of pecan and canola samples over the full spectral range. A summary of the best models for each sample set of individual oils is shown in Table 25. These results indicated that pretreatment via normalization and 9 point Savitzky-Golay (Gap 2) smoothing resulted in the best model obtained for pecan oil samples and the second best model for canola oil samples. Normalization with smoothing and first derivative pretreatment resulted in the best model obtained for canola oil

samples and the second best model for pecan oil samples. Either of these pretreatments accounted for all of the best pecan sample models as well as the three best canola sample models. Accordingly, each pretreatment was evaluated for all permutations of T0 and final week sample exclusions of both oils in the combined data set.

Table 25. Best individual oil peroxide value models for each sample set.

Q-Value	PLS Factors	Normalization (SNV)	Smoothing (9 Point S-G Gap 2)	Derivatization	Samples Excluded from Model	n	
0.5030	2	+	+	-	T0	18	Pecan
0.4648	2	+	+	1st	-	21	
0.4412	2	+	+	-	T0 and Week 14	15	
0.3936	2	+	+	-	Week 14	18	
0.4692	2	+	+	1st	T0	18	Canola
0.4560	3	+	+	-	Week 15	18	
0.4444	3	+	+	1st	-	21	
0.4308	2	+	-	-	T0 and Week 15	15	

Results of models from the sixteen sample permutations using normalization with smoothing and first derivative pretreatment are shown in Table 26 ranked by Q-value. Unlike the case for conjugated diene models, quality for peroxide models appeared higher when more samples were included. The sole exceptions were models which excluded only the T0 samples from either oil (see Table 26 bottom). Models were also generally worse than corresponding ones for conjugated dienes, which ranged from Q of 0.5145 to 0.6946 (Table 21). The few exceptions to this observation included the model made with all 42 samples (Q = 0.5913 for peroxides and 0.5440 for conjugated dienes).

Table 26. Combined data models of peroxide values using SNV and first derivative pretreatment ranked by Q-value.

Q-Value	PLS Factors	Pecan Samples Excluded		Canola Samples Excluded		n
		T0	Week 14	T0	Week 15	
0.5981	3	X		X		36
0.5922	3	X	X			36
0.5913	3					42
0.5878	3		X			39
0.5849	3		X		X	36
0.5829	3		X	X		36
0.5769	3			X	X	36
0.5761	3				X	39
0.5736	3	X	X		X	33
0.5625	3	X			X	36
0.5610	3	X	X	X		33
0.5516	3	X		X	X	33
0.5471	3	X	X	X	X	30
0.5376	3		X	X	X	33
0.5229	3			X		39
0.4667	4	X				39

Results of models from the sixteen sample permutations using normalization and smoothing pretreatment are shown in Table 27 ranked by Q-value. As was the case with conjugated diene models, the change in pretreatment shifted the order of sample sets relative to model quality. Without applying the first derivative, the trend of improved quality with additional samples was clarified.

All eight models comprising the bottom half of the quality rankings excluded the final week of canola samples (Table 27). The worst four models among those also excluded the T0 canola samples. The worst two of those four models excluded the T0 pecan samples. The worst of those two models also excluded the final week of pecan samples. This pattern repeated throughout all eight models which excluded the final week of canola samples, with the single caveat that the top two of those models had effectively identical Q-values.

Table 27. Combined data models of peroxide values using SNV and smoothing pretreatment ranked by Q-value.

Q-Value	PLS Factors	Pecan Samples Excluded		Canola Samples Excluded		n
		T0	Week 14	T0	Week 15	
0.6200	4			X		39
0.6155	4					42
0.6090	4		X			39
0.6081	4		X	X		36
0.6001	4	X		X		36
0.5964	5	X				39
0.5891	5	X	X			36
0.5851	4	X	X	X		33
0.5573	4		X		X	36
0.5573	4				X	39
0.5506	4	X			X	36
0.5423	4	X	X		X	33
0.5357	4			X	X	36
0.5332	4		X	X	X	33
0.5157	4	X		X	X	33
0.5042	4	X	X	X	X	30

Among the eight models comprising the top half of the quality rankings, the worst four models excluded the T0 pecan samples. The worst two of those four models also excluded the final week of pecan samples. Also among those four models, the worst and best excluded the T0 canola samples. This pattern repeated throughout all eight of the top models.

The best full spectrum peroxide model using the combined sample set was obtained with normalization and smoothing and afforded a Q-value of 0.6200. Only the initial canola oil samples were excluded, and model quality was better than those obtained from individual oils.

5.2.3 Predictive NIR Models for Lipid Oxidation using Selected Wavenumbers

Sample sets for the best combined models for conjugated dienes (oxidation index) and peroxides were analyzed to determine if they could be improved by eliminating wavenumbers from consideration. The e coefficient (Wu et al., 1995) provided information about how well the

oxidation parameters of interest aligned with overall changes among spectra of different oil samples during ASL testing. The cross validation regression coefficients t test (Martens & Martens, 2000) enabled selection of wavenumbers based on how critical they were to changes in the analyte of interest among ASL samples. Both approaches favor spectral regions which are not only relevant but also where repeated measurements of a sample are stable.

Analyses using the sample set for the best conjugated dienes (oxidation index) model [Q value 0.7214, normalization (SNV) and smoothing pretreatment, n = 33 excluding T0 canola oil samples and the final week samples of both oils] are reported below. However, problems were encountered using the sample set for the best peroxide value model [Q value 0.6200, normalization (SNV) and smoothing pretreatment, n = 39 excluding only the T0 canola oil samples]. Although numerous studies have ascribed bands to peroxides in the area of 4500 - 5000 cm^{-1} (see Appendix A), application of the cross validation regression coefficients t test to this sample set dropped this entire region from the model.

A comparison of the cross validation regression coefficient t test statistics for two models of peroxide values, one using the 39 samples from the data set providing the best peroxide value model and the other using the 33 samples from the data set providing the best conjugated diene model is shown in Figure 46. These sample sets differed only by the final week samples of each oil. In the 4500 - 5000 cm^{-1} region, a large difference between sample sets was observed in the t test values (Figure 46, top) which was largely attributable to an increased standard deviation (Figure 46, bottom) when the final week samples were included. While including these samples improved model quality for full spectrum models, it also increased variability in this critical spectral region that shifted the wavenumber selection process away from a region known for peroxide signals. When the t test was applied using the smaller sample set, wavenumbers in this critical region were identified as important. Therefore, the smaller set of 33 samples was used for wavenumber selection in peroxide value models as well as those for conjugated dienes.

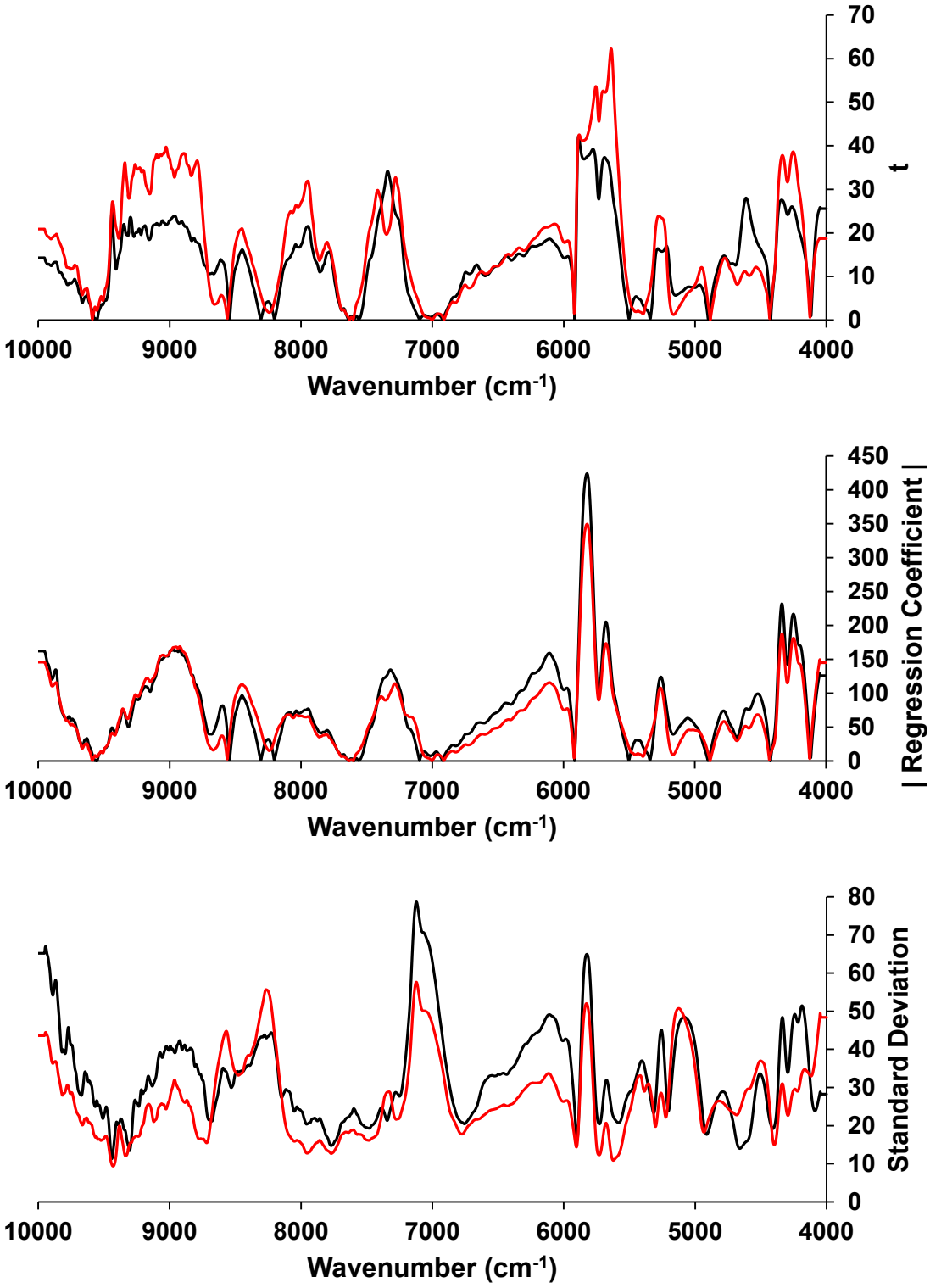


Figure 46. Comparison of regression coefficient statistics for two initial models of peroxide values made from different sample sets (— n = 39; — n = 33): t test values (top), absolute value (center) and standard deviation among cross validation submodels (bottom).

5.2.4 NIR Model Improvement: Wavenumber Selection using the e Coefficient

The value of e, the ratio of differences among groups to those within groups, was calculated for the combined set of 33 samples with pretreatments that provided the highest-quality model, namely, normalization (standard normal variate) followed by smoothing (9 point Savitzky-Golay Gap 2). The spectrum of e values is presented in Figure 47a. Three large peaks were evident at 5164 cm^{-1}, 5688 cm^{-1} and 5828 cm^{-1}. The first of these likely corresponded to the rice flour matrix and water as a hydroxyl combination band has been attributed to polysaccharides at 5181 cm^{-1} and multiple sources cite bands associated with moisture from 5155 to 5200 cm^{-1} (see Appendix A). The middle peak likely corresponded to CH stretching from methylene groups noted at 5682 cm^{-1} in a number of lipid oxidation references and at 5666 cm^{-1} and 5675 cm^{-1} in aliphatic compounds. The final peak also likely arose from methylene groups noted at 5797-5829 cm^{-1} in lipid oxidation references as well as at 5797 cm^{-1} in aliphatic compounds.

The 5164 cm^{-1} peak could indicate differences among samples in the relative composition of oil to flour as measured. The e value was calculated by comparing variation among mean absorbances for each sample batch (numerator) with variation among individual tablets prepared from the same oil (denominator). If measurements for each batch did not reflect the true mean value of the nominal composition (7.5% oil), model quality would suffer. Even with correct batch preparation, inaccuracies could arise from failure of the mean absorbance of the fifteen tablets measured for each batch to converge on the true mean.

This peak could also indicate differences in moisture content among samples. The forty-two original batches were measured on fourteen different days over the course of nearly four months. The flour matrix used could have taken up different amounts of moisture based on the time of exposure of samples during batch preparation as well as the relative humidity in the lab. Fluctuating moisture content among samples would also have impaired model quality. Finally, the peak at 5164 cm^{-1} could arise at least in part from lipid hydroperoxide decomposition induced when oxidized oil was mixed with rice flour, placing it in contact with protein amine groups

(Schaich & Karel, 1976), followed by protonation from antioxidants present, by lipids, or by protein donor groups.

Values of e for all wavenumbers were within a single order of magnitude (Figure 47a). Thus, only limited improvements to the model were possible by excluding less valuable wavenumbers because no regions within the spectrum were overwhelmingly better than any others. While extremely low value wavenumbers possessing significant noise and little useful information could be excluded, removal of those with a moderate amount of useful information would degrade the quality of the model much more rapidly than in the case where certain wavenumbers were favored by extremely large discrepancies in e values.

Further analysis was made of the individual components of e (Figure 47b). The numerator of e indicated discrepancies among sample batches at 7116 cm^{-1} as well as 8264 cm^{-1} which were degraded in the final calculation of e due to within group variation in those areas. The former is coincident with hydroxyl vibrations indicative of water, alcohols and other hydroxylated species, while the latter is a -CH overtone region common in oils. Also, the numerator indicated a much broader peak above 5000 cm^{-1} than was evident from e (Figure 47a) which was degraded by within group variation, particularly at 5028 cm^{-1} and 5272 cm^{-1} which coincides with starch from the rice flour.

These observations indicate compositional variation as the relative content of oil and rice flour varied within the samples. Additionally, moisture contents may have fluctuated both within and among sample batches. Both of these phenomena could have impaired model quality.

5.2.4.1 Conjugated Diene Model Improvement Based on e Values

Attempts were made to improve model quality by eliminating wavenumber ranges with e values below a cutoff threshold. Figure 48 shows the Q-value for the best model obtained for the full spectrum (e cutoff 0.0) and at wavenumber ranges determined by e cutoffs at intervals of 0.25. As shown in Figure 49, as the e cutoff applied increased, wavenumber ranges used in models were split, truncated, or lost entirely.

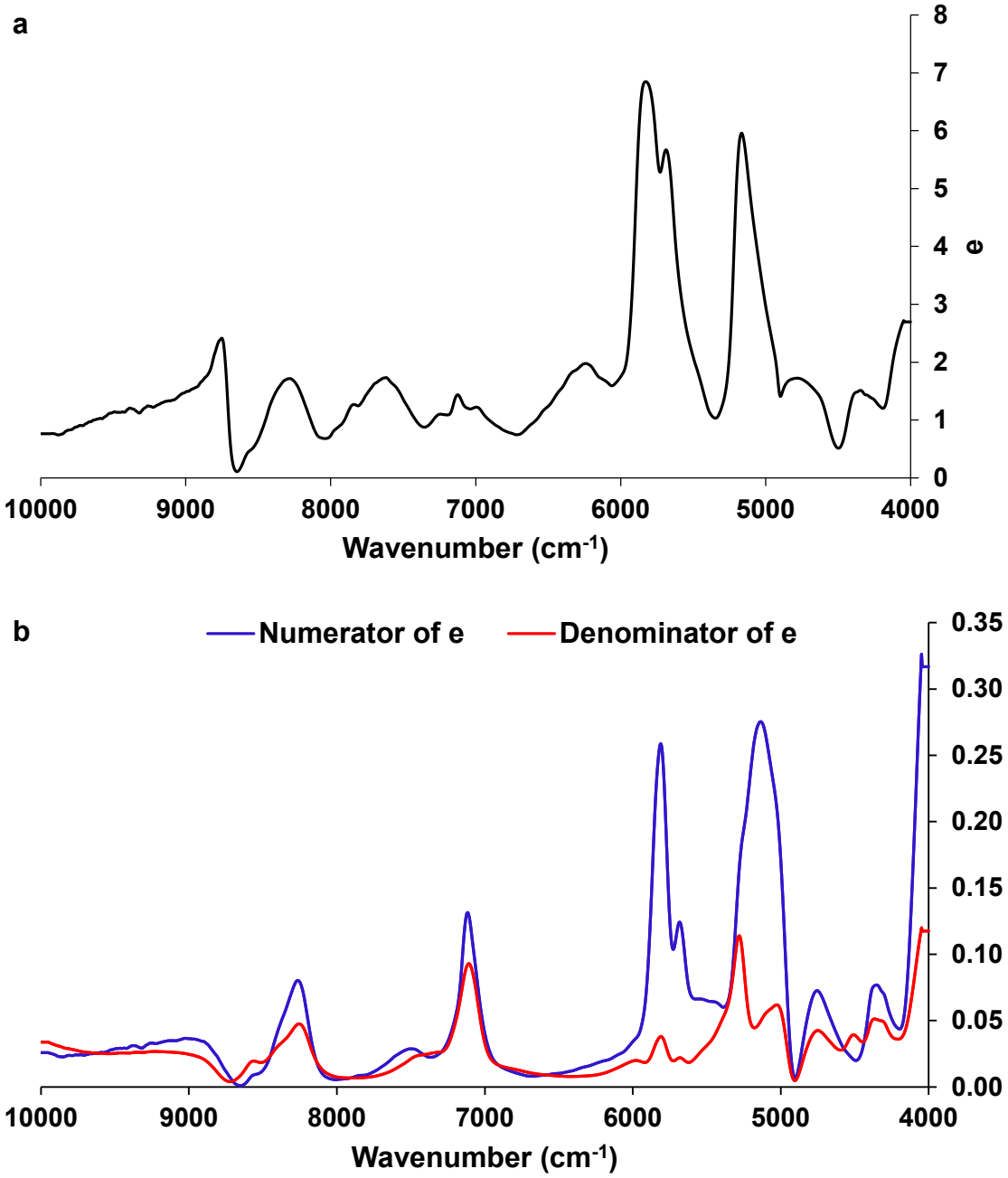


Figure 47. Plots of (a) e coefficient values; and (b) the numerator and denominator of e.

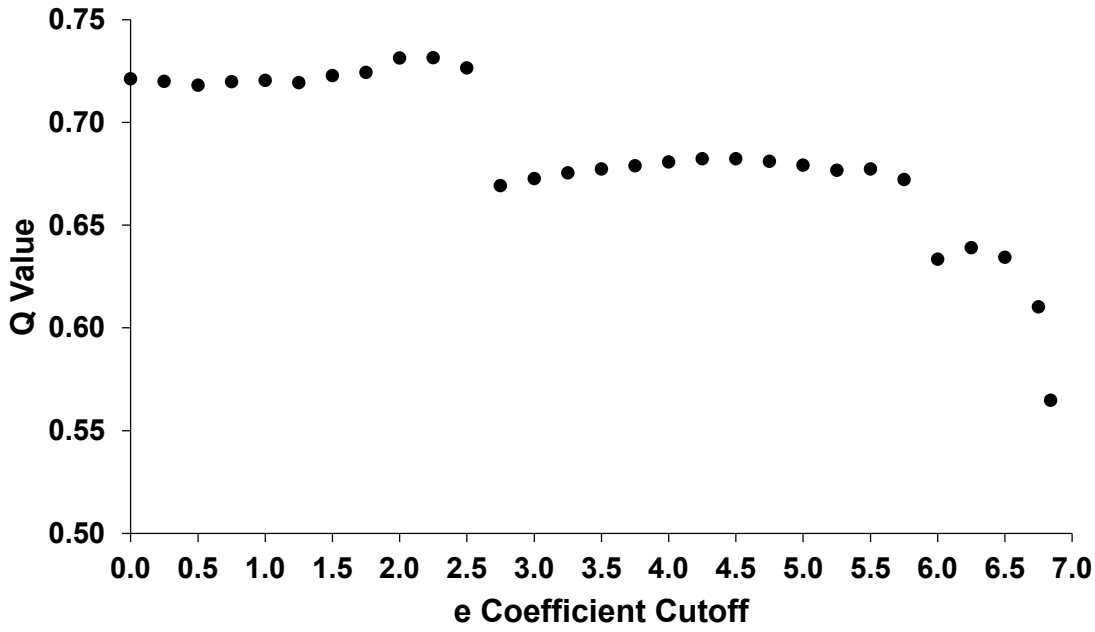


Figure 48. Refinement of conjugated dienes model by wavenumber selection based on e.

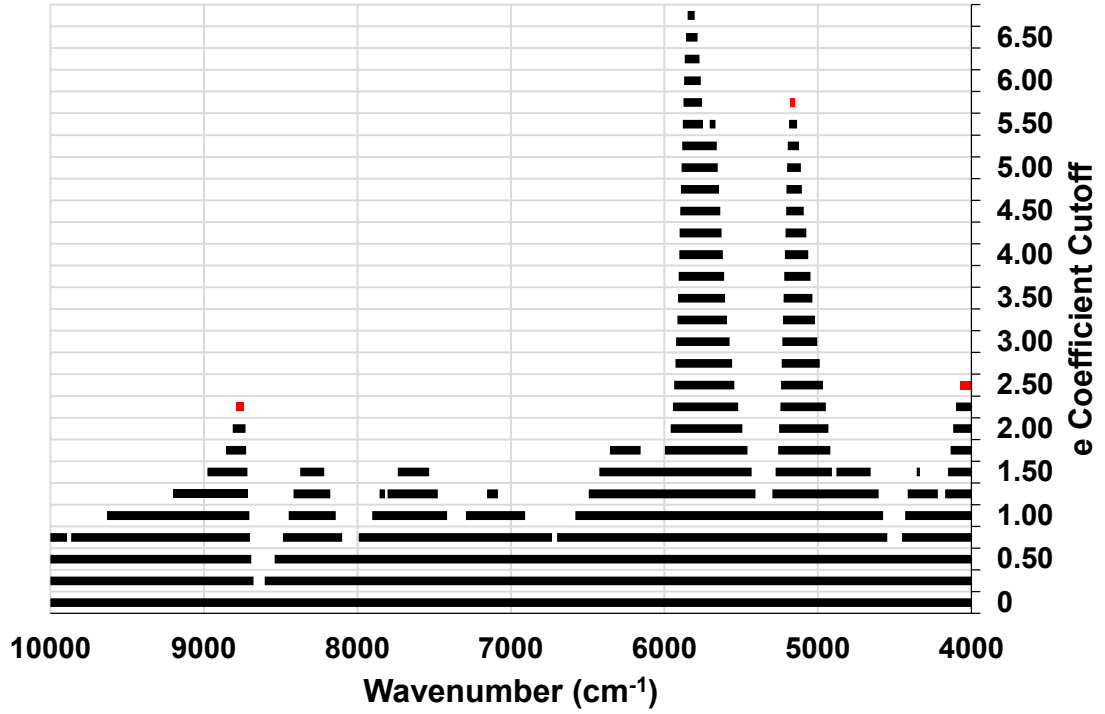


Figure 49. Wavenumber ranges used in models after application of e coefficient cutoffs.

Improvements were slight, while a minor drop in quality was observed from an e cutoff of 2.25 to 2.5 and sudden drops in quality were observed at two points (2.5 to 2.75 and 5.75 to 6.0). These breakpoints corresponded to the loss of entire regions shown in red in Figure 49. The slight drop in model quality corresponded to the loss of the 8740-8788 cm^{-1} region, which falls within the 8547-8873 cm-1 range for the second overtone of the CH stretch band. The first large drop coincided with the loss of 4000-4076 cm^{-1}, a region known to contain bands for a methylene CH combination in lipids (4049 cm-1) as well as CH and CC stretch combination bands (4000 cm^{-1} and 4063 cm^{-1}) in starch. The second large drop occurred upon loss of the 5148-5180 cm^{-1} region atop the first of the three largest peaks in the e value spectrum. As discussed above, this range encompasses bands for moisture and its combination with hydroxyls in starch from the flour matrix. The highest Q-value (0.7336) was obtained by applying a cutoff of 2.5 with reinstatement of the 8740-8788 cm^{-1} region.

5.2.4.2 Conjugated Diene Model Improvement Based on the Numerator of e

Attempts to improve model quality were also made using the numerator of e to determine the importance of ranges exhibiting differences between groups but with e values decreased due to within group variation. Figure 50 shows the Q-value for the best model obtained for the full spectrum and at wavenumber ranges (Figure 51) determined by cutoff intervals of 0.025.

Elimination of wavenumbers at lower cutoff values initially detracted from the quality of the full spectrum model. Q-values returned to those of the initial model using a cutoff of 0.15, the first point at which the 7076-7148 cm^{-1} range (red in Figure 51) was fully discarded. This corresponds to hydroxyl signals from water, alcohols and other oxygenated compounds, including hydroperoxides. A slight improvement in model quality peaked at a cutoff of 0.20 (Q-value 0.7288), corresponding to the removal of the shoulder from 5244-5272 cm^{-1} overlaying an area of increased within group variability (Figure 47b and Figure 51). This region is near bands reported for water and includes bands reported for starch, indicating that variation in moisture content or

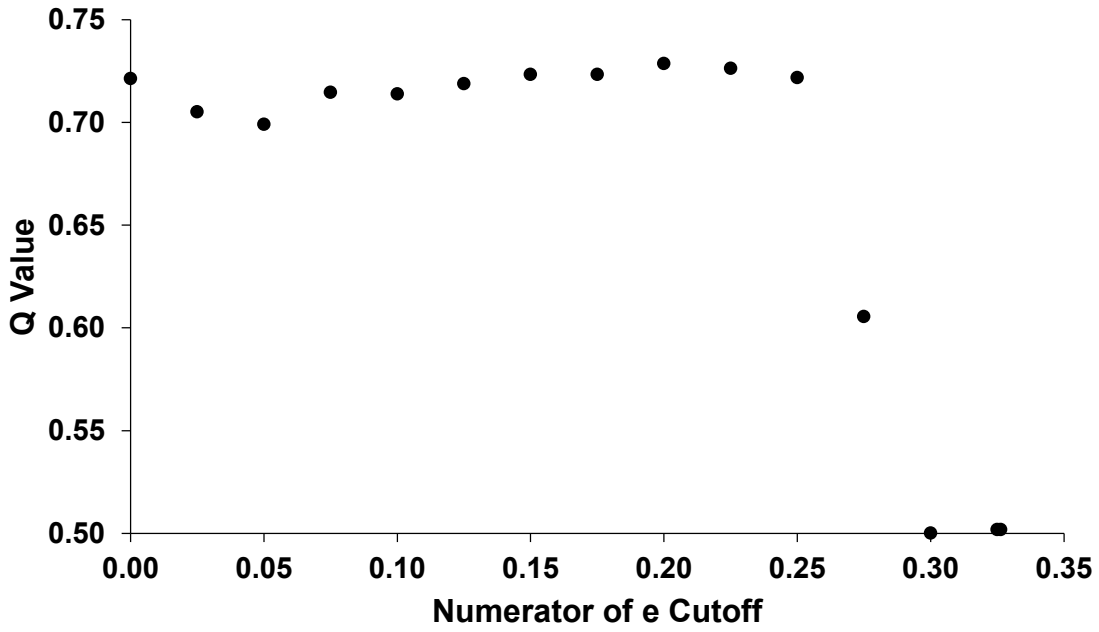


Figure 50. Refinement of conjugated dienes model by selection of wavenumbers based on the numerator of e.

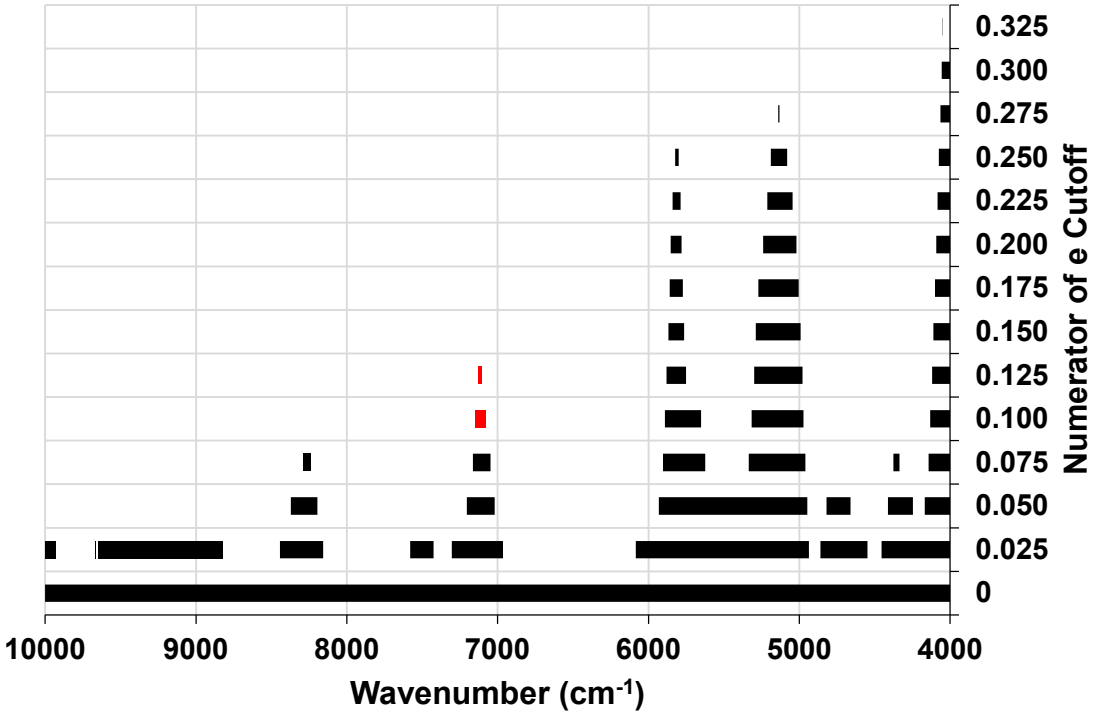


Figure 51. Wavenumber ranges used in models after application of numerator of e cutoffs.

composition (% oil) of the samples likely confounded the analysis. The three ranges in the best model included 4000-4092 cm^{-1}, 5020-5240 cm^{-1} and 5780-5852 cm^{-1}. Due to CH bands in the first and last regions (Appendix A), they may be linked to measurement of oxidation while the middle region likely accounted for variation in oil and possibly moisture content of the samples.

5.2.4.3 Peroxide Value Model Improvement Based on e Values

The value of e for the combined set of 33 samples was also used to assess peroxide value model quality. The same pretreatments that provided the best conjugated dienes model, namely, normalization (standard normal variate) followed by smoothing (9 point Savitzky-Golay gap 2), were also found to provide the best peroxide value models. Attempts were made to improve model quality by eliminating wavenumber ranges with e values below a cutoff threshold. Figure 52 shows Q-values for the best models obtained for the full spectrum (e cutoff 0.0) and at wavenumber ranges determined by e cutoffs at intervals of 0.25.

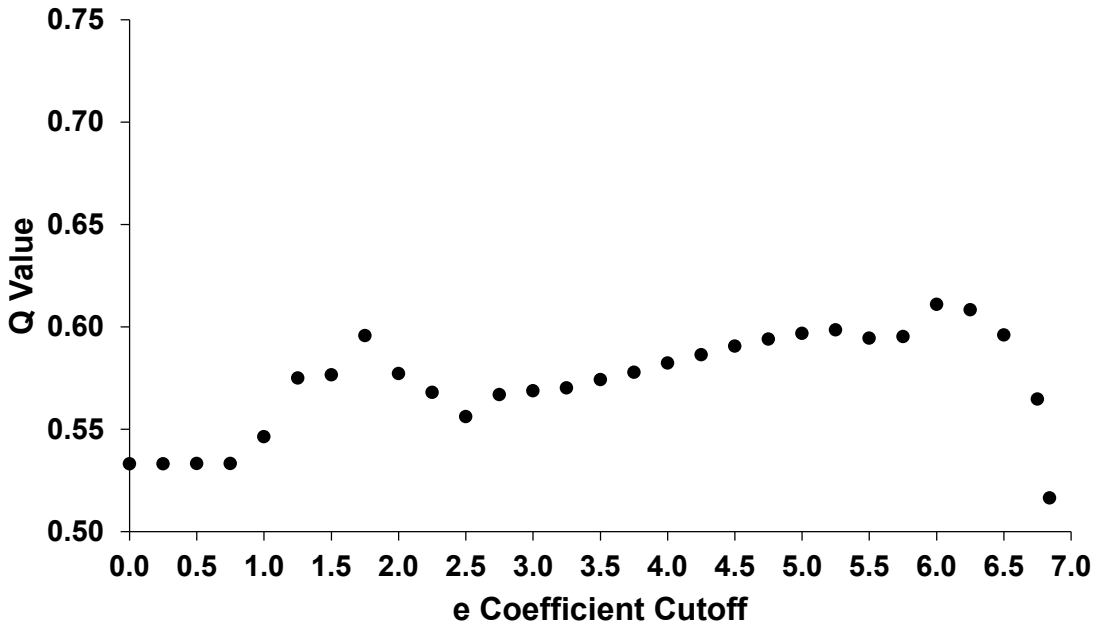


Figure 52. Refinement of peroxide value model by wavenumber selection based on e.

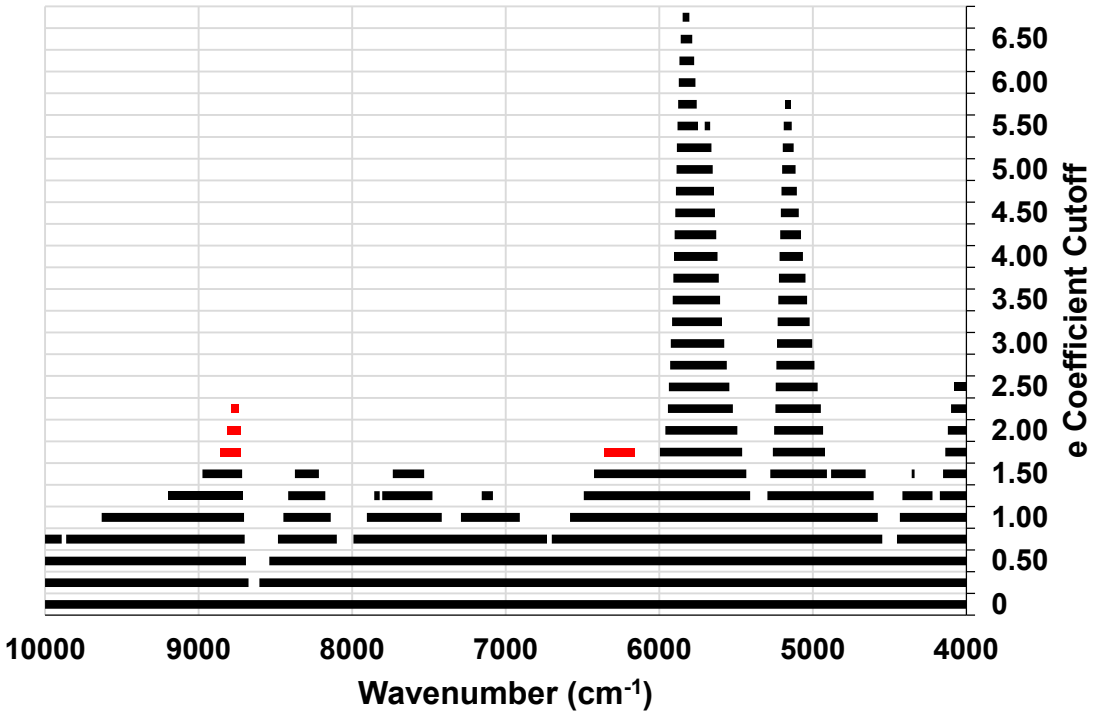


Figure 53. Wavenumber ranges used in models after application of e coefficient cutoffs.

Q values for peroxide models were lower than those for conjugated dienes, indicating a drop in model quality for this analyte. Unlike those for conjugated dienes, peroxide models steadily improved as wavenumbers were removed from consideration, with the exception of quality losses from a peak at an e cutoff of 1.75 to 2.50. The wavenumber ranges lost in this drop include 6156-6356 cm^{-1} and 8724-8856 cm^{-1}, both shown in red in Figure 53. The former range includes CH bands for alkenes as well as hydroxyl bands of water or alcohols. The latter is the same region observed in quality losses in the conjugated dienes model and corresponds to the second overtone of the CH stretch band (8547-8873 cm^{-1}).

The best Q-value (0.6110) was observed at an e cutoff of 6.0, in which only the narrow region from the largest peak in the spectrum of e values (5764 - 5872 cm^{-1}) was retained. The model was improved by combining that region with 6156-6356 cm^{-1} and 8724-8856 cm^{-1}, resulting in a Q-value of 0.6250. When the six samples dropped from the best peroxide value model were reinstated, model quality improved to a Q-value of 0.6636.

5.2.4.4 Peroxide Value Model Improvement Based on the Numerator of e

Attempts to improve model quality were also made using the numerator of e to determine the importance of the ranges exhibiting differences between groups but minimized in the e calculation due to within group variation. Figure 54 shows Q-values for the best models obtained for the full spectrum and at wavenumber ranges determined by e numerator cutoffs of 0.025. Figure 55 shows the wavenumber ranges meeting the respective cutoffs.

Elimination of wavenumbers at lower cutoff values initially improved model quality to a maximum Q-value of 0.5725 at a cutoff of 0.075. The subsequent drop in model quality corresponded to the loss of 4336-4376 cm^{-1} and 8240-8292 cm^{-1}, shown in red in Figure 55. The former region corresponds to numerous reports of CH bands from lipid oxidation studies and is close to signals for combination bands for CH stretching and methylene deformation as well as OH and CC stretching in starch. Lipid oxidation studies have also reported CH bands in the latter region (Kaddour et al., 2006; Cozzolino et al., 2005; Armenta et al., 2007).

All of the NIR studies we found for peroxides indicated the importance of wavenumbers in the 4500-5000 cm^{-1} range (Appendix A). As the 4660-4820 cm^{-1} range was dropped in going from a cutoff of 0.050 to 0.075, a refined analysis was performed to determine if this region should have been discarded. Q values are shown in Figure 56 and wavenumber ranges in Figure 57. The analysis indicated the highest quality model occurred at a cutoff of 0.065 (Q-value 0.5748). This model relied upon seven wavenumber ranges, 4000-4152 cm^{-1}, 4288-4396 cm^{-1}, 4712-4792 cm^{-1}, 4956-5352 cm^{-1}, 5500-5912 cm^{-1}, 7036-7176 cm^{-1}, and 8220-8324 cm^{-1}, which included an area near wavenumbers attributed to peroxides in several NIR studies of lipid oxidation (Appendix A).

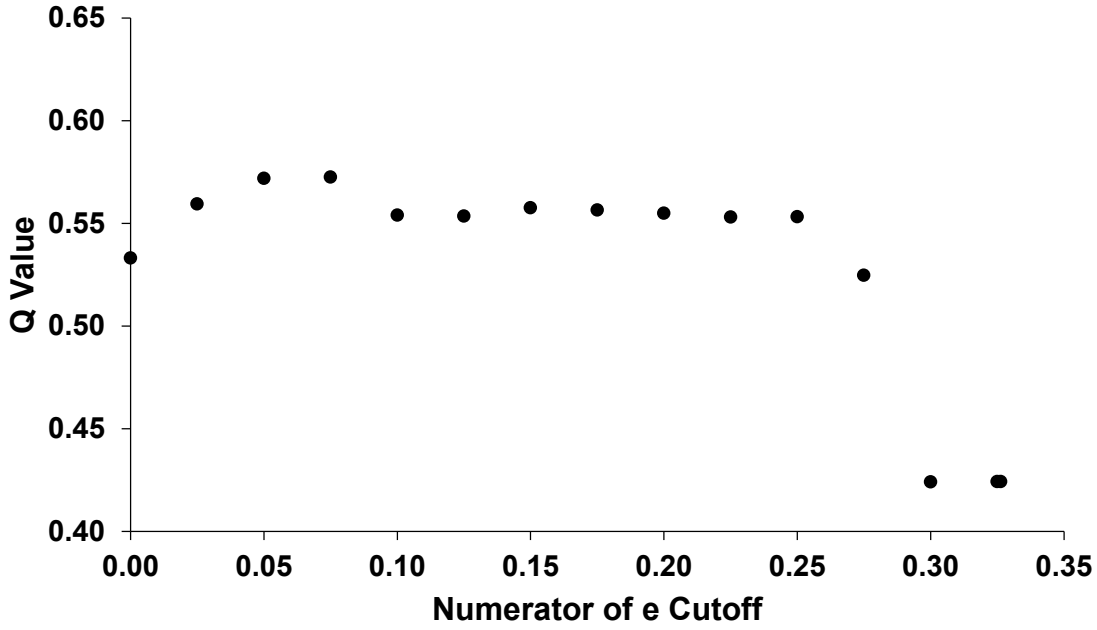


Figure 54. Refinement of peroxide value model by selection of wavenumbers based on the numerator of e.

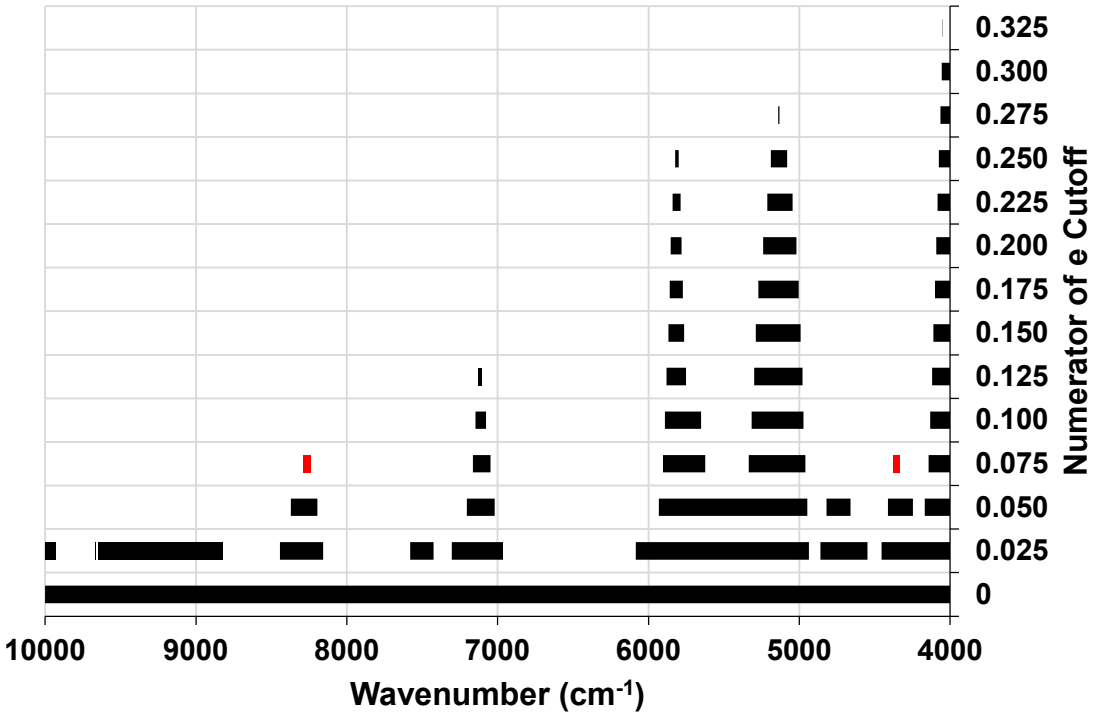


Figure 55. Wavenumber ranges used in models after application of numerator of e cutoffs.

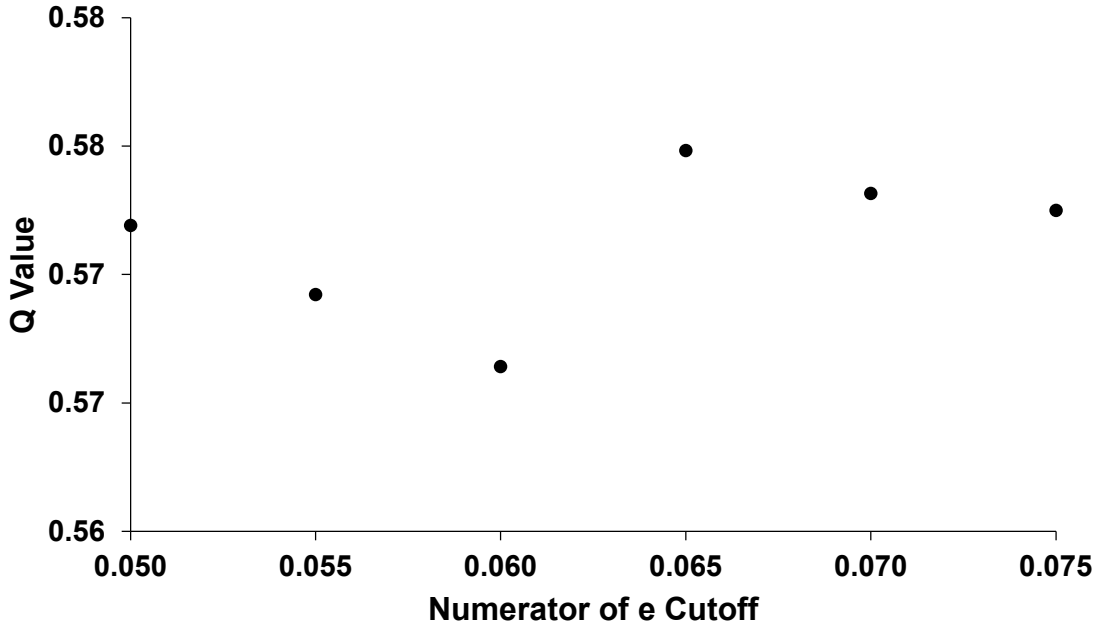


Figure 56. Q-values for models from refined numerator of e cutoffs.

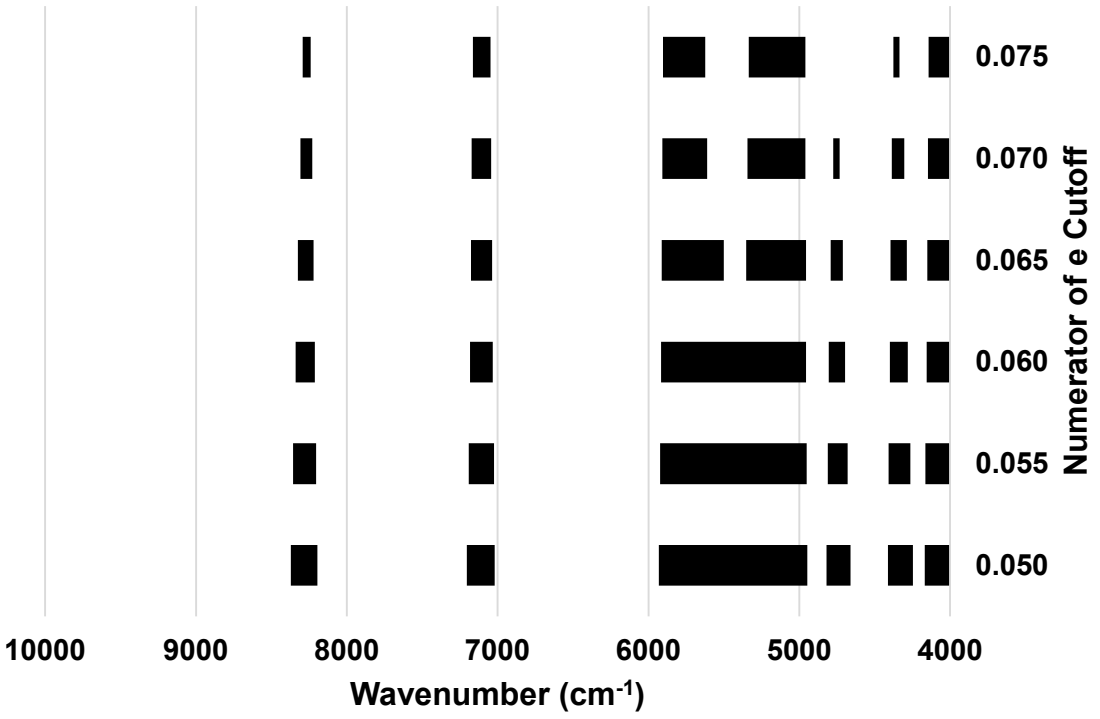


Figure 57. Wavenumber ranges for refined numerator of e cutoffs.

5.2.5 NIR Model Improvement: Cross Validation Regression Coefficients t Test

Attempts to improve model quality using cross validation regression coefficients t test values were made by two approaches. First, wavenumbers with t values failing to pass significance testing indicating they were non zero at a given level of confidence were eliminated. Second, similar to e value testing, a cutoff ranking approach was used to find the best models. Each approach was also iterated according to the jackknife procedure (Martens & Martens, 2000; Westad & Martens, 2000). Regression coefficient statistics for the best model for conjugated dienes from the combined set of 33 samples with normalization and smoothing pretreatments are shown in Figure 58. This full spectrum model afforded a Q-value of 0.7214. The corresponding statistics for the peroxide value model using the same sample set and pretreatments are shown in Figure 59. This full spectrum model afforded a Q-value of 0.5332, well below the Q-value of 0.6200 for the best model of peroxide values.

5.2.5.1 Conjugated Diene Model Improvement by Significance Testing of t

As shown in Figure 58, the overwhelming majority of wavenumbers exhibited regression coefficients with t test values well above zero. Thus, a high level of confidence was used to provide the best opportunity for model improvement during iterations of significance testing. A confidence level of 99.5% was selected, which at 32 degrees of freedom provided a t test critical value of 2.740. Wavenumber selection converged within two iterations as shown in Table 28. Improvements were modest as most wavenumbers passed testing and were retained (Figure 60). Thus, a ranking process was investigated to attempt further improvement.

Table 28. Conjugated diene models from significance testing of cross validation regression coefficient t test values at 99.5% confidence.

Model	Q Value	Wavenumbers Used
Initial (Full Spectrum)	0.7214	1501
Iteration 1	0.7287	1291
Iteration 2	0.7293	1267
Iteration 3	"	"

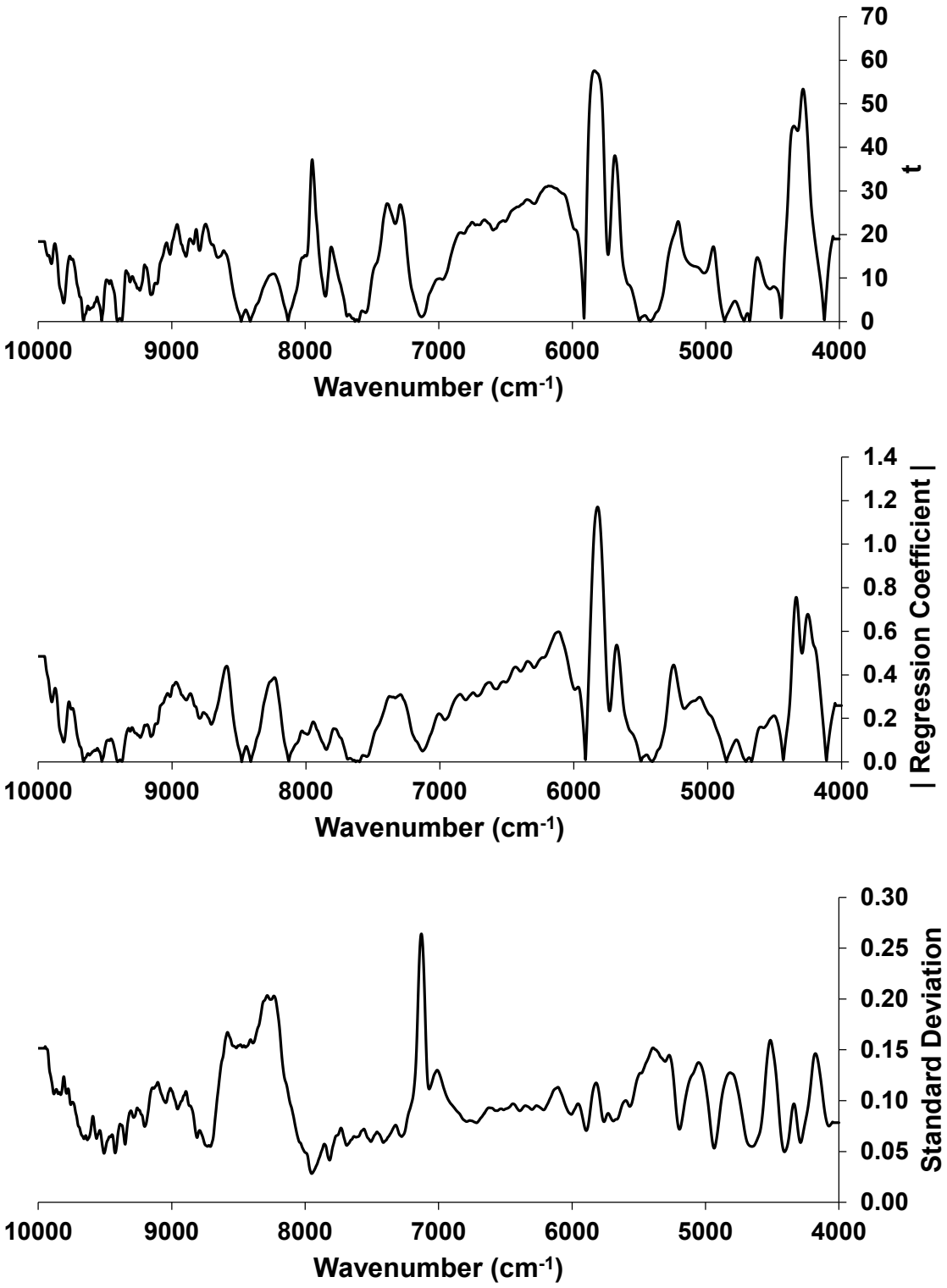


Figure 58. Regression coefficient statistics for best initial conjugated dienes model: t test values (top), absolute value (center) and standard deviation among cross validation submodels (bottom).

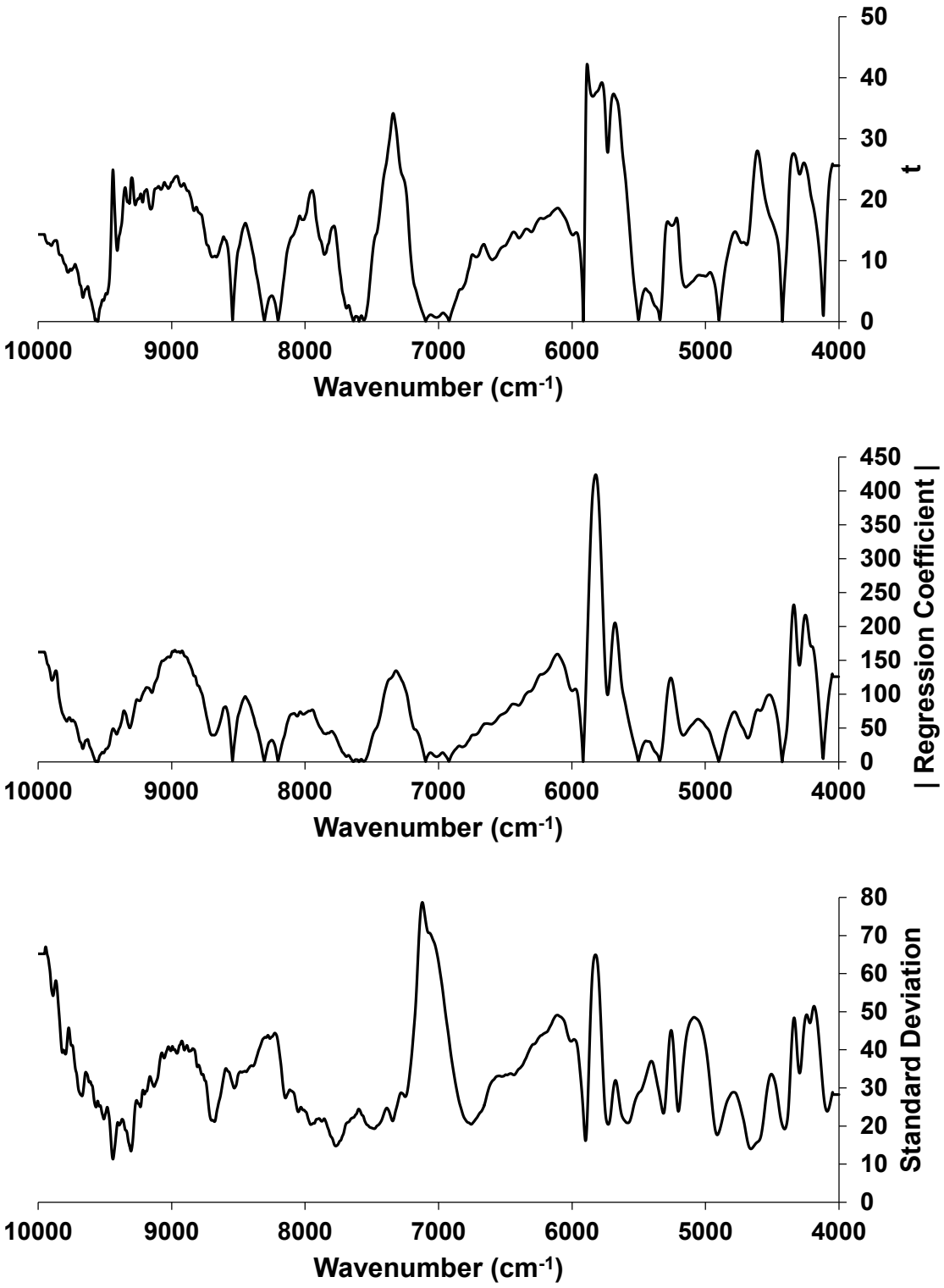


Figure 59. Regression coefficient statistics for initial peroxide values model subjected to improvement procedures: t test values (top), absolute value (center) and standard deviation among cross validation submodels (bottom).

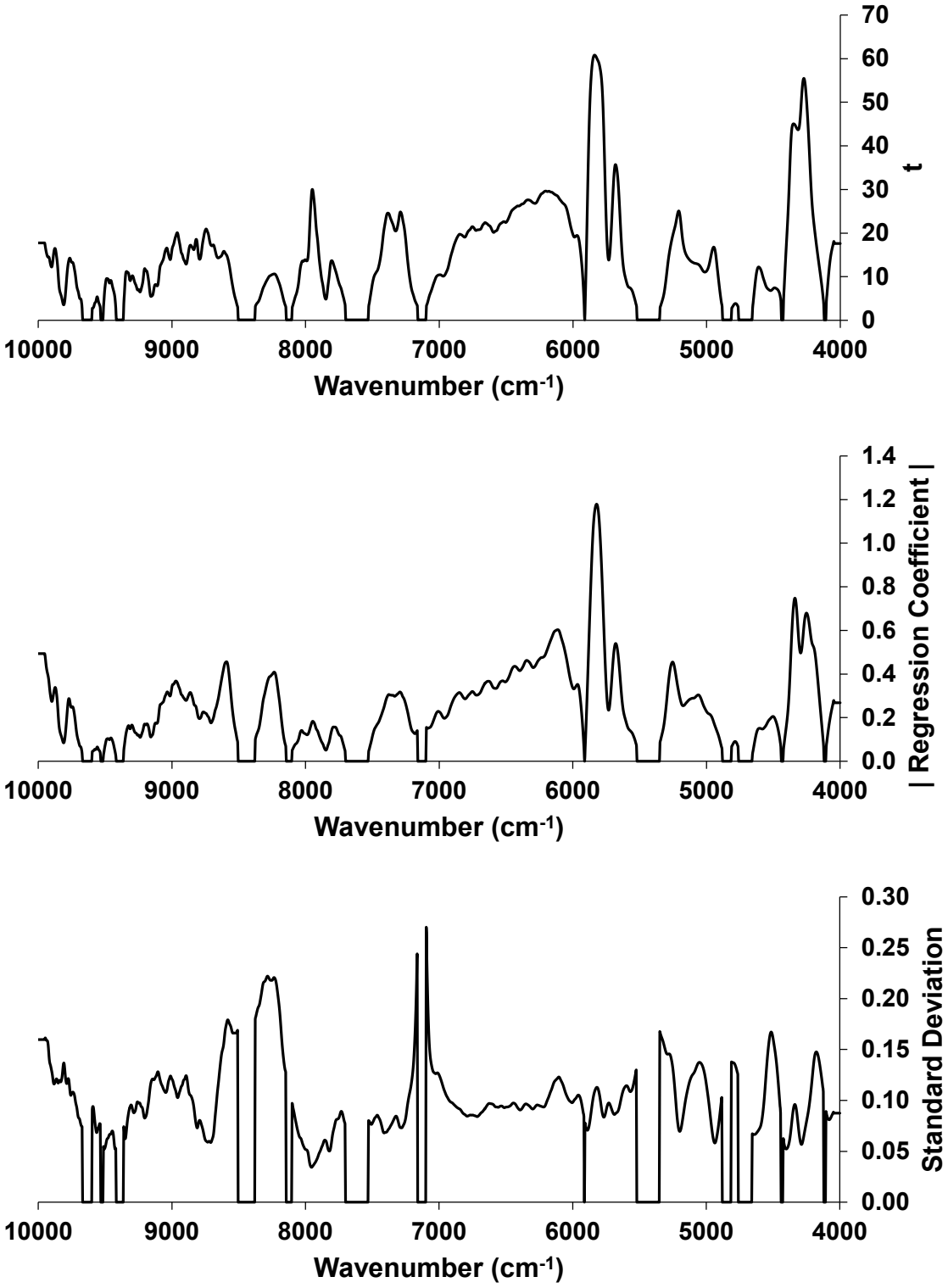


Figure 60. Regression coefficient statistics for conjugated dienes model after second iteration of significance testing (99.5% Confidence): t test values (top), absolute value (center) and standard deviation among cross validation submodels (bottom).

5.2.5.2 Conjugated Diene Model Improvement by Ranking of t

Because model optimization stalled during significance testing due to the large size of t values relative to the critical value in the t-distribution for the number of samples used, ranking was also evaluated. During each iteration, a series of models was assessed using wavenumbers with t values at or above a threshold cutoff value. The model with the highest Q-value was selected and t values were recalculated using regression coefficients of the overall model as well as the standard deviation of the regression coefficients of submodels generated during cross validation. These values were used in a cutoff analysis for the following iteration. Q-values for models from these cutoff analyses beginning with the best initial model for conjugated dienes (see values of t in Figure 58) are shown in Figure 61, while statistics for the best models are shown in Table 29.

Table 29. Conjugated diene models from ranking of cross validation regression coefficient t test values.

Model	Q Value	Wavenumbers Used	t Cutoff (Best Model)
Initial (Full Spectrum)	0.7214	1501	10
Iteration 1	0.7389	911	7.5
Iteration 2	0.7409	887	0

Application of a t cutoff of 10 to the data from the initial model removed over 1/3 of the wavenumbers and improved the Q-value to 0.7389. A modest increase in Q-value to 0.7409 was observed using a t cutoff of 7.5 on data from the first iteration model. Cutoff analysis on the second iteration model (Figure 61 bottom) failed to generate any improvement, causing wavenumber ranges to converge and ending the optimization. The ranking procedure resulted in a slight improvement in model quality beyond significance testing (Q = 0.7409 versus 0.7293) due to the exclusion of data from additional wavenumbers evident from comparison of Figure 60 and Figure 62. Although sixteen wavenumber ranges were isolated, empirical testing showed that use of only eight of those further improved model quality to a Q-value of 0.7519 (see Table 32).

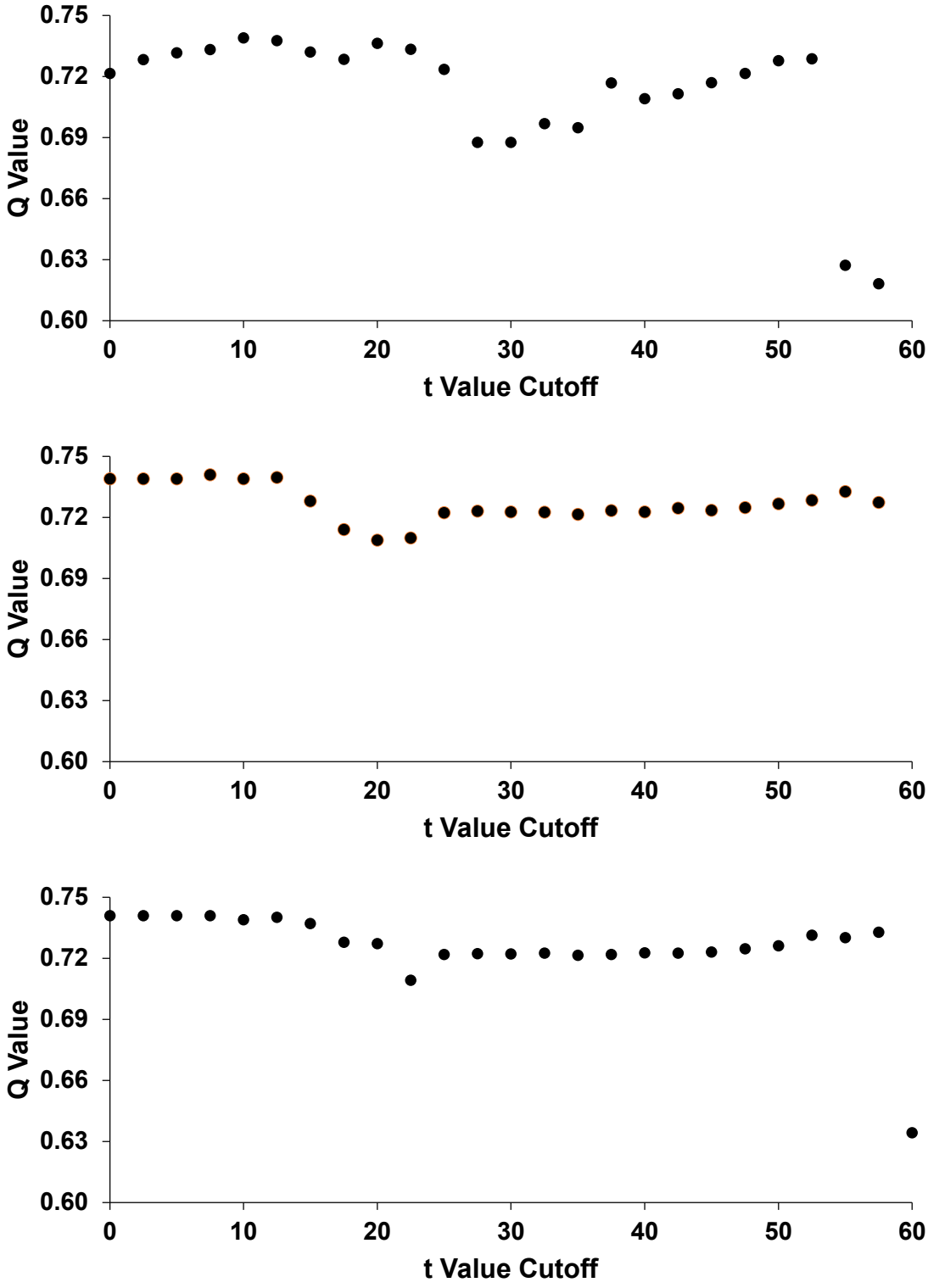


Figure 61. Conjugated dienes model Q-values: wavenumber selection based on cross validation regression coefficients t test cutoff thresholds for the best initial model (top), best first iteration model (middle) and best second iteration model (bottom).

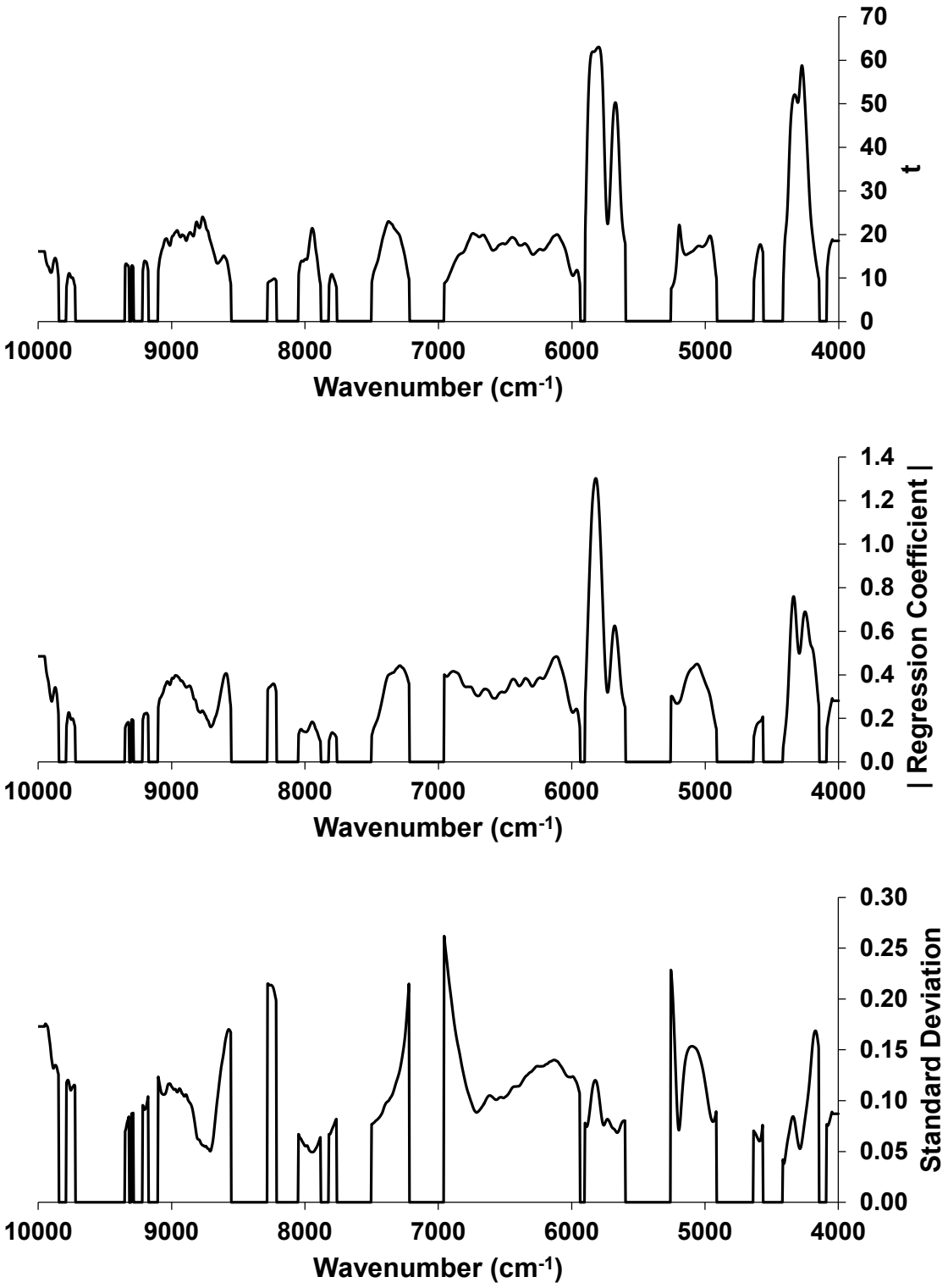


Figure 62. Regression coefficient statistics for conjugated dienes model after second iteration of regression coefficient t test ranking: t test values (top), absolute value (center) and standard deviation among cross validation submodels (bottom).

5.2.5.3 Peroxide Value Model Improvement by Significance Testing of t

As was the case with conjugated dienes, the overwhelming majority of wavenumbers for the initial peroxide value model (Figure 59) exhibited regression coefficients with t test values well above zero. A high level of confidence (99.5%) was used to provide the best opportunity for model improvement during iterations of significance testing. The t test critical value was 2.740. Optimization converged to a final set of 1255 wavenumbers within four iterations, resulting in a very modest improvement in model quality (Table 30).

Table 30. Peroxide value models from significance testing of cross validation regression coefficient t test values at 99.5% confidence.

Model	Q Value	Wavenumbers Used
Initial (Full Spectrum)	0.5332	1501
Iteration 1	0.5486	1322
Iteration 2	0.5562	1263
Iteration 3	0.5565	1256
Iteration 4	0.5567	1255
Iteration 5	"	"

Significance testing from the peroxide value model resulted in slightly fewer wavenumbers than for the conjugated dienes model. Notable changes included a much broader gap between wavenumbers used around key hydroxyl bands near 7000 cm^{-1} in the former [6788 - 7196 cm^{-1} (Figure 63)] than in the latter [7100 - 7162 cm^{-1} (Figure 60)]. Also, the peroxide model appeared to include slightly more wavenumbers in the 4500 - 5000 cm^{-1} range often cited in the lipid oxidation literature.

A modest quality increase resulted upon application of the 1255 wavenumbers isolated by significance testing to the set of 39 samples which provided the best peroxide value model. As the Q-value rose from 0.6200 for the full spectrum model to only 0.6360 using wavenumbers from significance testing, ranking procedures were investigated to determine if the model could be further improved.

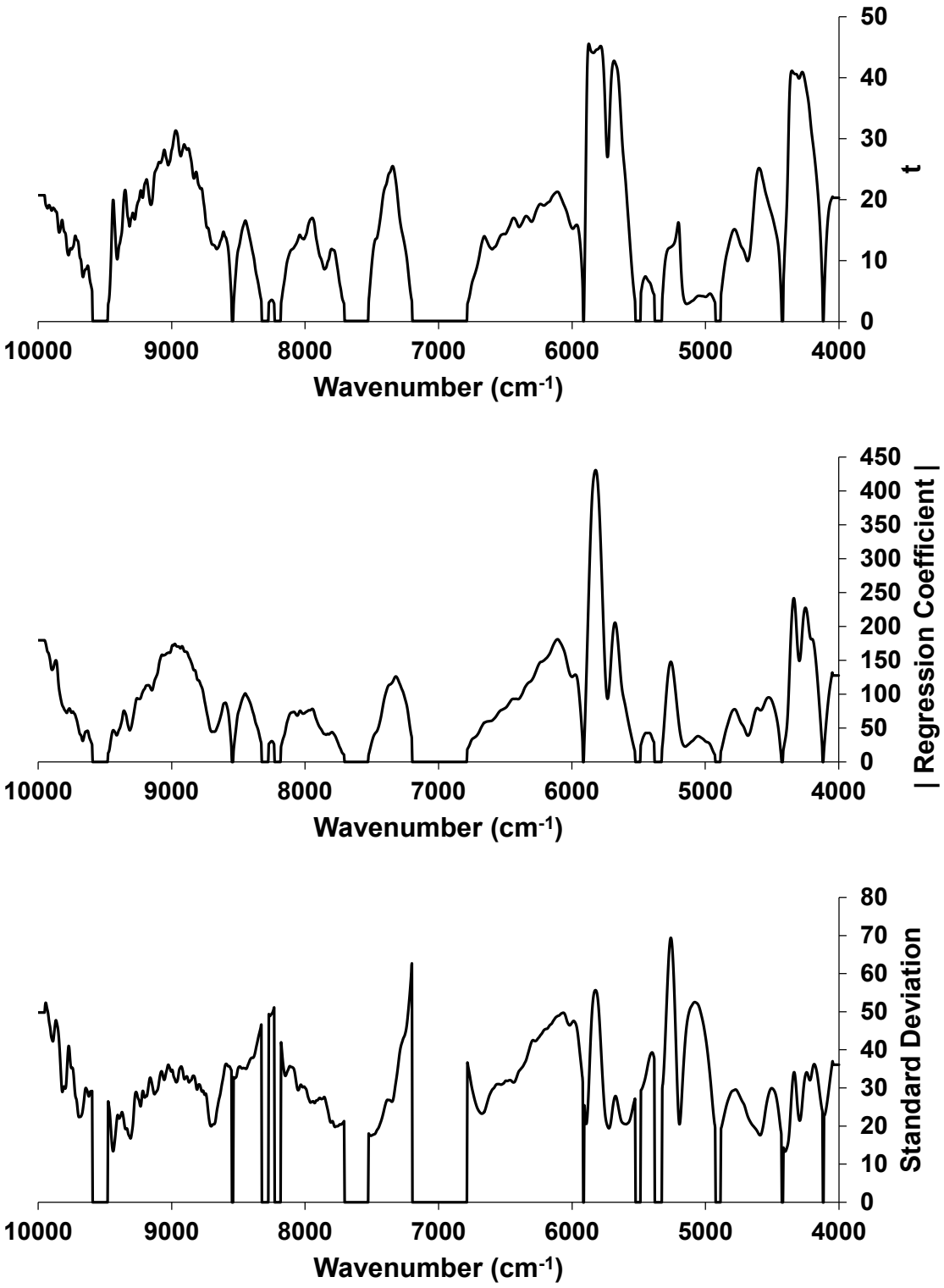


Figure 63. Regression coefficient statistics for peroxide value model after fourth iteration of significance testing (99.5% confidence): t test values (top), absolute value (center) and standard deviation among cross validation submodels (bottom).

5.2.5.4 Peroxide Value Model Improvement by Ranking of t

The ranking procedure (see 5.2.5.2) was also used on the peroxide value model from the combined data set of 33 pecan and canola oil samples. Q-values for models from cutoff analyses beginning with the initial model (Figure 59) are shown in Figure 64, while statistics for the best models are shown in Table 31.

Table 31. Peroxide value models from ranking of cross validation regression coefficient t test values

Model	Q Value	Wavenumbers Used	t Cutoff (Best Model)
Initial (Full Spectrum)	0.5332	1501	12
Iteration 1	0.5763	864	24
Iteration 2	0.6453	226	34
Iteration 3	0.6467	178	0

Analysis of the data from the initial full spectrum model indicated that a t cutoff of 12 resulted in the first maximum in Q-value of 0.5763 (Figure 64 top) associated with the elimination of over 42% of the spectrum. A t cutoff of 24 was applied in the next iteration (Figure 64 center), eliminating 85% of the spectrum and improving quality to a Q-value of 0.6453. In the final iteration applied (Figure 64 bottom), a t cutoff of 34 resulted in a slight improvement in Q-value to 0.6467. This model was unable to be further improved by the ranking procedure and retained only 178 wavenumbers or 12% of the full spectrum. Unlike the conjugated dienes model resulting from the ranking procedure (Figure 62), the peroxide value model (Figure 65) included many fewer wavenumbers in only five discrete regions of the NIR spectral range. This result reflected the smaller number of hydroxyl bands relative to -CH bands in NIR. Wavenumbers isolated included 4176 - 4360 cm^{-1}, 4592 - 4632 cm^{-1}, 5216 - 5260 cm^{-1}, 5768 - 5884 cm^{-1}, and 8808 - 9116 cm^{-1}.

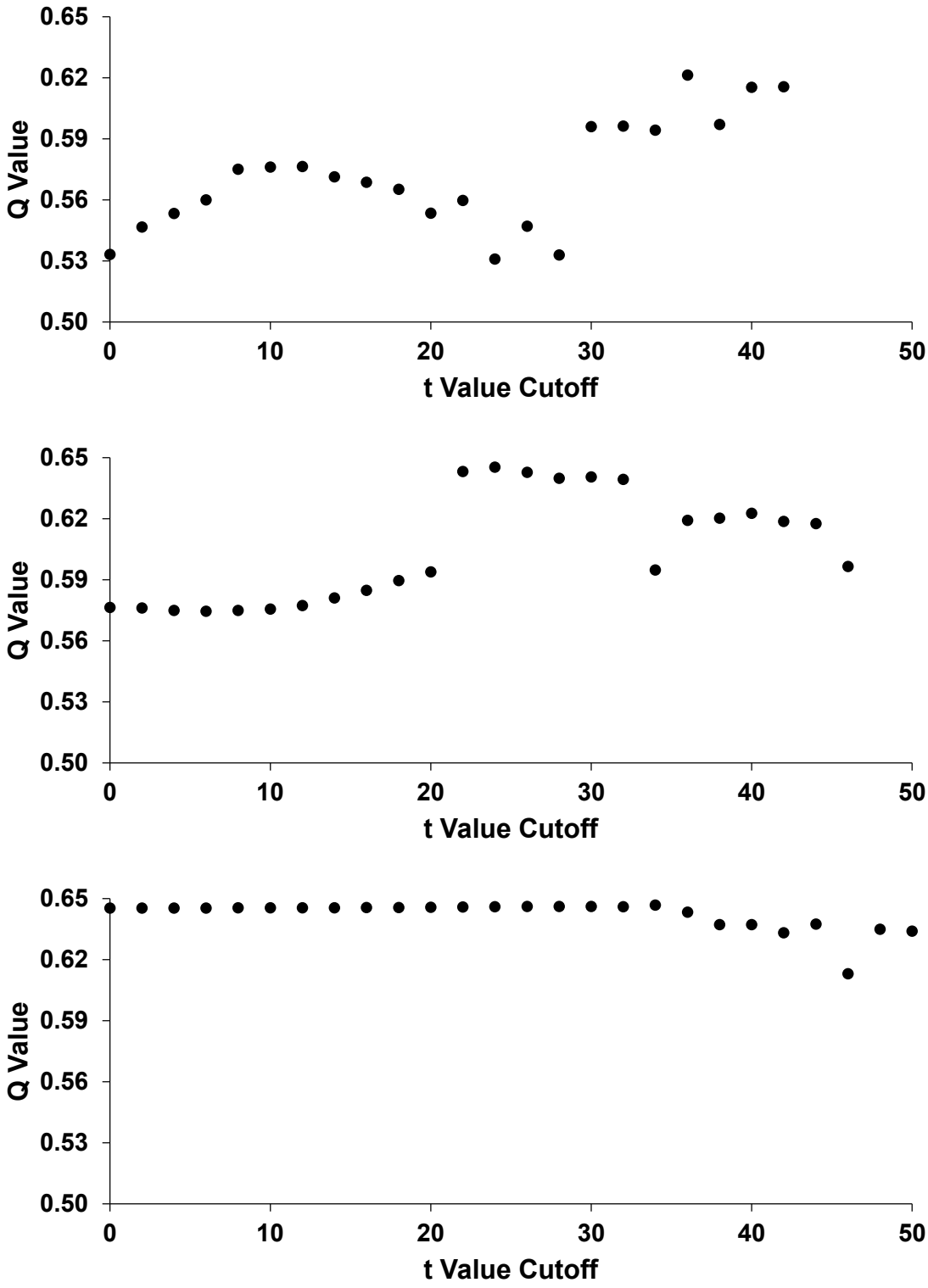


Figure 64. Peroxide value model Q-values: wavenumber selection based on cross validation regression coefficients t test cutoff thresholds for the initial model (top), best first iteration model (center) and best second iteration model (bottom).

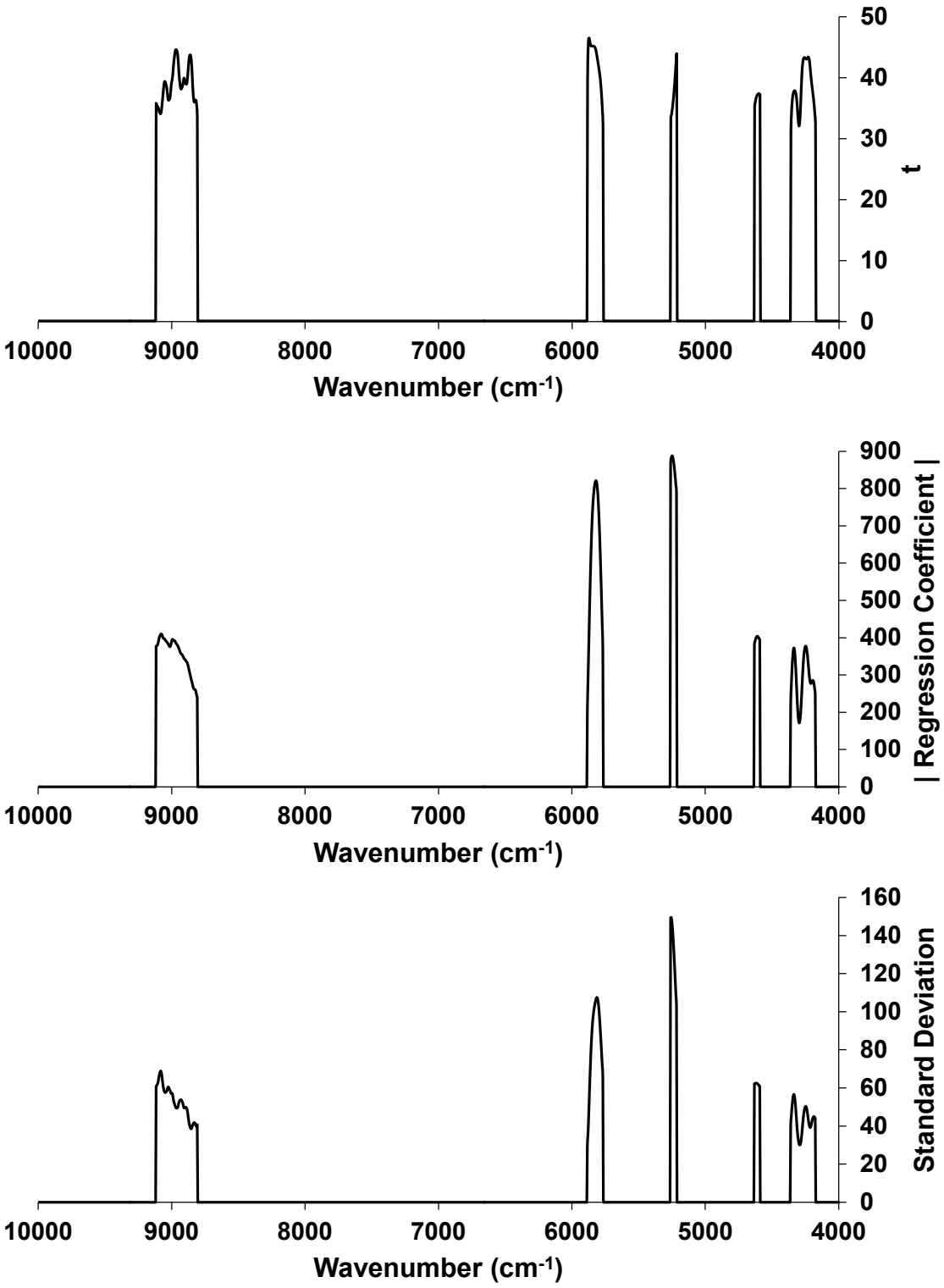


Figure 65. Regression coefficient statistics for best peroxide value model after third iteration of regression coefficient t test ranking: t test values (top), absolute value (center) and standard deviation among cross validation submodels (bottom).

The 178 wavenumbers isolated from the ranking procedure were then applied to create a model using the combined oil sample set (n = 39 with only canola T0 samples excluded) that provided the optimum full spectrum peroxide value model. The result was a significant increase in model quality from a Q-value of 0.6200 for the full spectrum model to 0.7328 for the limited wavenumber model. It is known that outliers should be reevaluated after limiting wavenumbers since the irregularities in those samples may be tied to the spectral regions removed by the improvement procedure (Davies, 2001). In this instance, optimization of the best full spectrum model resulted in the loss of all wavenumbers around a critical spectral region where peroxides have been reported (4500 - 5000 cm^{-1}, see Appendix A). Although the model was not optimized directly, the use of a smaller sample set succeeded in isolating wavenumbers of interest that could then be applied to the larger sample set to improve the model. This underscores the importance of presenting a truly representative sample to the NIR. Not only does variation directly affect model quality, it also affects the optimization process and impacts the effectiveness of model improvement by wavenumber selection.

5.2.6 Summary of Best Models

Models using wavenumber ranges isolated by the various improvement techniques used on full spectrum models of the combined data set of both pecan and canola samples are presented in Figure 66 for conjugated dienes and Figure 67 for peroxide values. For both analytes, the numerator of e provided the least improvement, likely because variation among sample replicates was not taken into account. This effect translated to these models despite the fact that each relied solely upon a single average spectrum for each sample batch. Ranking of the cross validation regression coefficients t test values provided the best improvement in both cases. For conjugated dienes an empirical selection of ranges from those isolated by ranking provided further improvement. Wavenumbers from the top model for each analyte were applied to pecan, canola and combined sample data sets to generate the quantitative results assessed below.

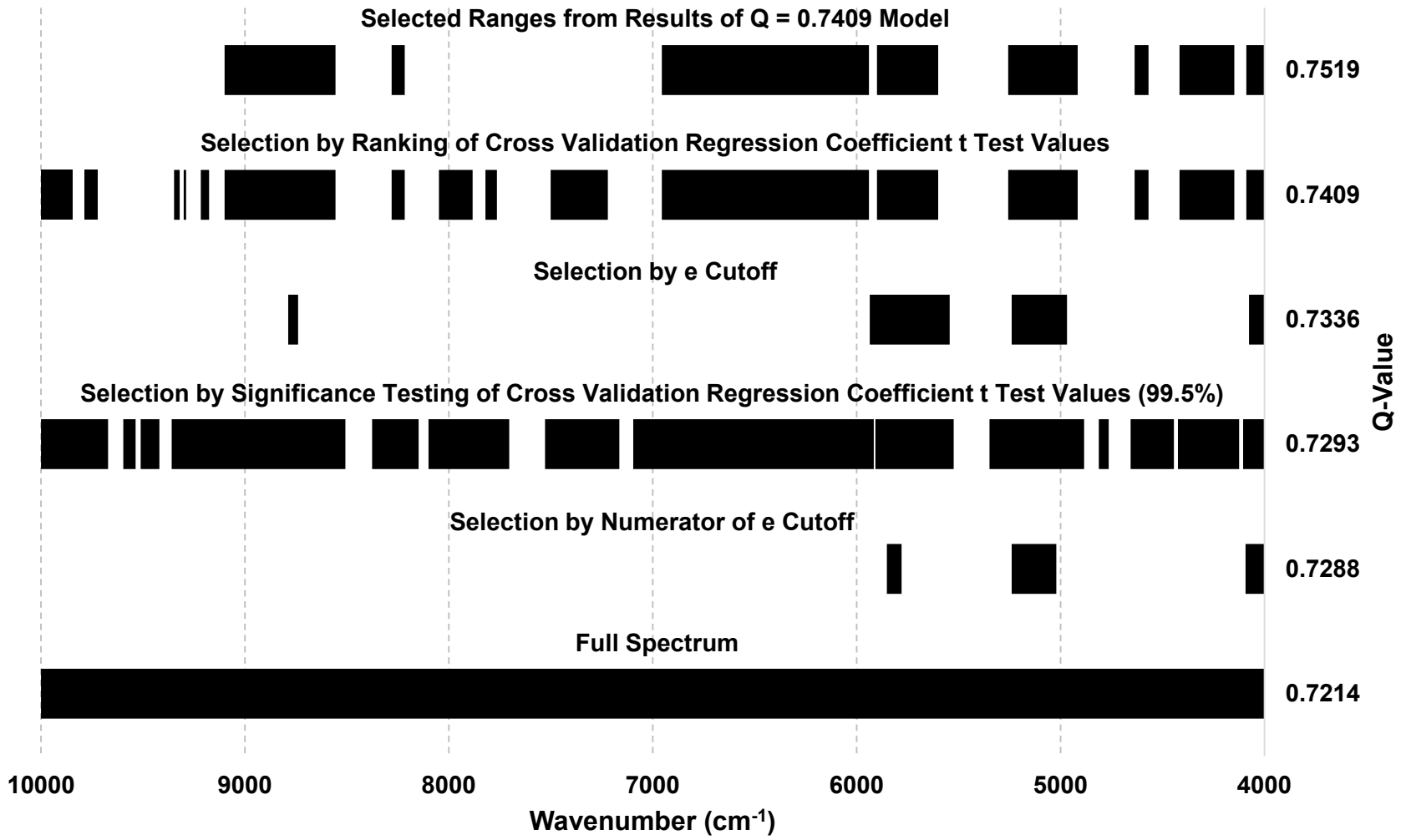


Figure 66. Summary of wavenumber ranges isolated by techniques to improve models of conjugated dienes.

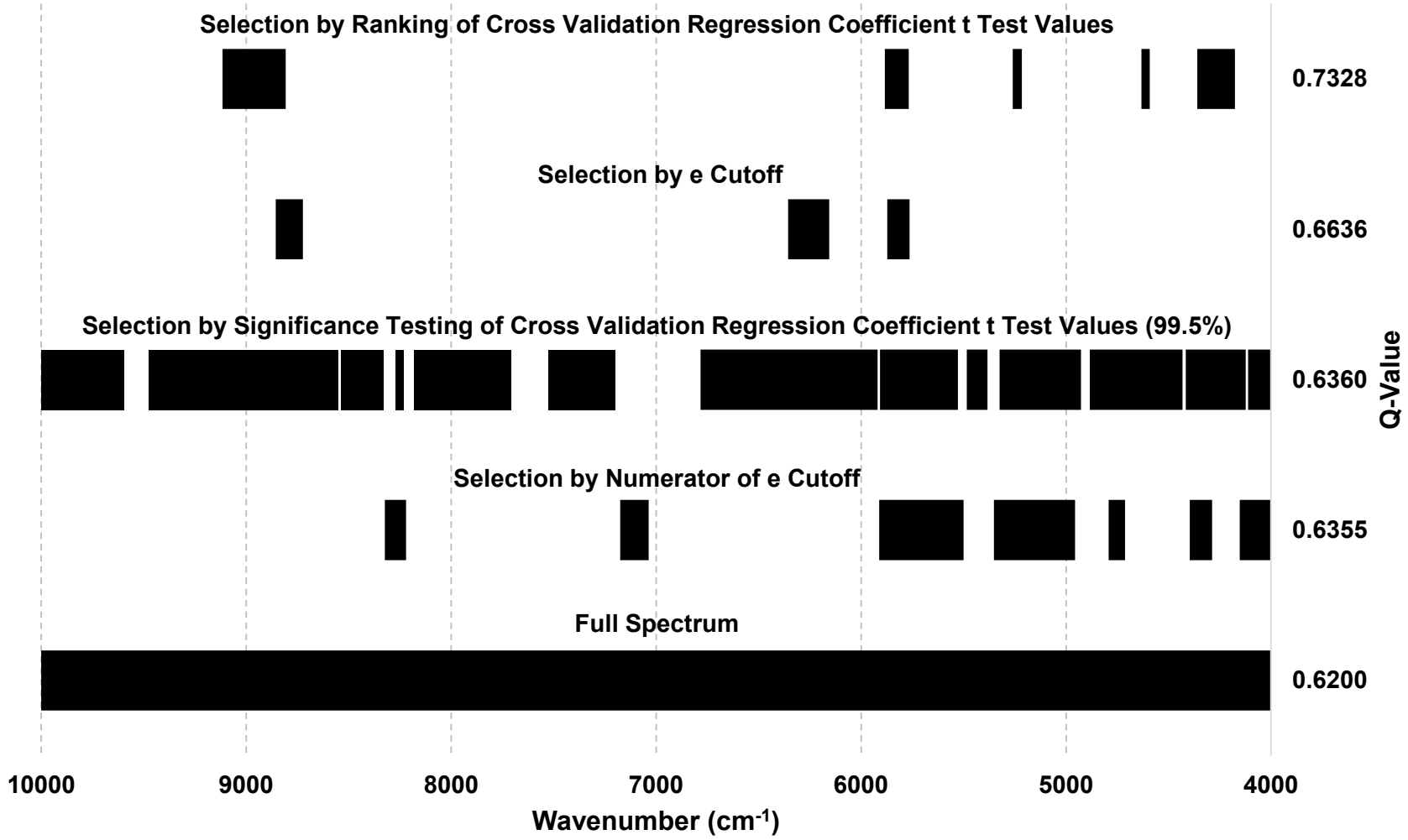


Figure 67. Summary of wavenumber ranges isolated by techniques to improve models of peroxide values.

5.2.6.1 Conjugated Dienes

Statistics of the best models for conjugated dienes are summarized in Table 32. Models shown were based on the full spectral range scanned by the NIR as well as selected wavenumber ranges isolated by the ranking procedure and a subset of those ranges empirically determined to improve model quality. This subset excluded wavenumbers in the 7000 - 8000 cm^{-1} region as well as all wavenumbers above 9100 cm^{-1}. The former range includes bands known to be associated with moisture near its low end while the latter is inherently variable as noise increases in FT-NIR spectra approaching 10,000 cm^{-1}.

Plots comparing original and NIR predicted values of conjugated dienes expressed as the oxidation index for the best models for each sample set are shown in Figure 68. Poor linearity was observed among samples made with pecan oil (Figure 68, top), while good linearity was observed for canola oil samples (Figure 68, center) and to a lesser extent for the combined samples (Figure 68, bottom). The disparity in fit was likely due to changes in the degree of oxidation of the respective samples.

Despite the large difference in peroxide values observed for the respective oils during the ASL, the oxidation index values for the two exhibited a substantial degree of overlap. This indicated that models for both oils should be of similar quality, while the opposite was observed. Although the increased maximum oxidation index for canola oil (1.3 vs. 1.0 for pecan oil) certainly contributed to the disparity in quality, closer inspection of the data revealed the respective distribution of oxidation index values in the samples as the likely explanation for the poor quality of the pecan oil sample model. While twelve of the fifteen batches (80%) used in the canola oil sample model had oxidation index values of at least 0.72, twelve of the eighteen batches (67%) used in the pecan oil sample model had oxidation index values of 0.60 or less.

Table 32. Statistics of best models for conjugated dienes.

		Calibration		Validation				
Wavenumbers	Q-Value	r^2	SEC	r^2	SECV	BIAS	$SD_{Reference}$	RPD
Pecan Samples								
(Normalization (SNV), Smoothing and 1st Derivative)							All Samples Except T0 (n = 18)	
Full Spectrum (4000 - 10,000 cm^{-1})	0.6266	0.6710	0.1523	0.5739	0.1799	-0.006836	0.2655	1.476
Selected Ranges *	0.6309	0.6831	0.1495	0.5855	0.1773	-0.006675	0.2655	1.497
Selected Ranges **	0.6331	0.6808	0.1500	0.5870	0.1771	-0.006533	0.2655	1.499
Canola Samples								
(Normalization (SNV) and Smoothing)							All Samples Except T0 and Final Week (n = 15)	
Full Spectrum (4000 - 10,000 cm^{-1})	0.7252	0.9817	0.04469	0.8858	0.1437	0.007160	0.3305	2.299
Selected Ranges *	0.7510	0.9727	0.05462	0.8945	0.1299	0.008679	0.3305	2.545
Selected Ranges **	0.7721	0.9743	0.05296	0.9099	0.1206	0.008042	0.3305	2.741
Pecan and Canola Samples								
(Normalization (SNV) and Smoothing)							All Samples Except Pecan Final Week and Canola T0 and Final Week (n = 33)	
Full Spectrum (4000 - 10,000 cm^{-1})	0.7214	0.8687	0.1225	0.7824	0.1694	0.003700	0.3380	1.996
Selected Ranges *	0.7409	0.8610	0.1260	0.7931	0.1621	0.002298	0.3380	2.086
Selected Ranges **	0.7519	0.8630	0.1251	0.8027	0.1585	0.001175	0.3380	2.133

* 4000 - 4088 cm^{-1}, 4148 - 4416 cm^{-1}, 4568 - 4636 cm^{-1}, 4916 - 5256 cm^{-1}, 5600 - 5900 cm^{-1}, 5940 - 6956 cm^{-1}, 7220 - 7500 cm^{-1}, 7764 - 7820 cm^{-1}, 7884 - 8048 cm^{-1}, 8216 - 8280 cm^{-1}, 8556 - 9100 cm^{-1}, 9176 - 9216 cm^{-1}, 9288 - 9300 cm^{-1}, 9320 - 9348 cm^{-1}, 9724 - 9788 cm^{-1} and 9848 - 10,000 cm^{-1}

** 4000 - 4088 cm^{-1}, 4148 - 4416 cm^{-1}, 4568 - 4636 cm^{-1}, 4916 - 5256 cm^{-1}, 5600 - 5900 cm^{-1}, 5940 - 6956 cm^{-1}, 8216 - 8280 cm^{-1}, and 8556 - 9100 cm^{-1}

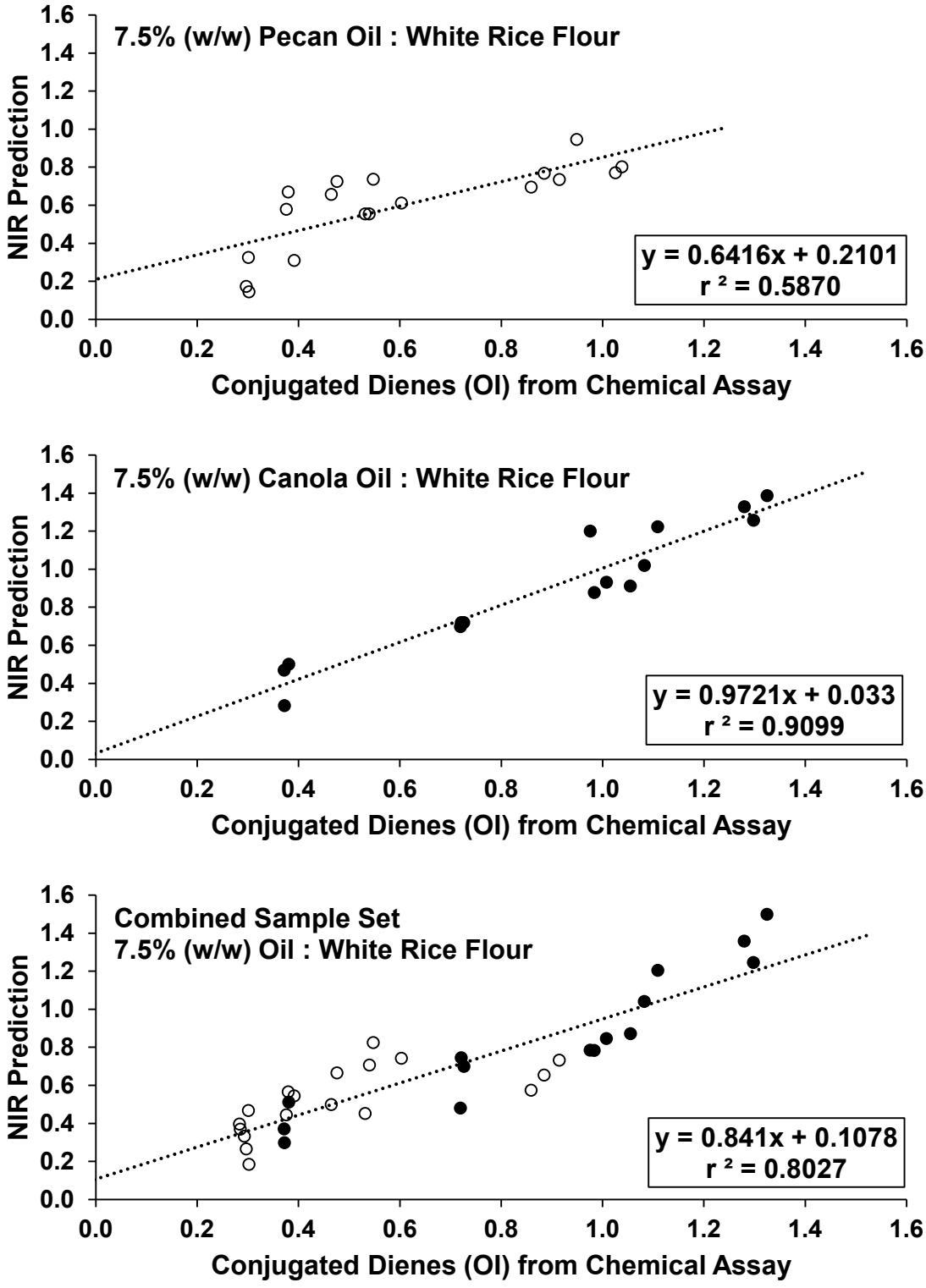


Figure 68. Original vs. NIR predicted values of conjugated dienes (oxidation index) for best models of pecan oil samples (n = 18) (top), canola oil samples (n = 15) (center) and combined sample set (n = 33) (bottom): Pecan oil samples (○); Canola oil samples (●).

A key sample presentation issue was whether the amount of sample scanned by the miniaturized sample system was sufficiently representative to provide a true average spectrum of the material. The data in Table 32 indicate that the values of SECV were much larger than those of the BIAS for each model. The former is a measure of precision while the latter is a measure of accuracy (NIRCal 5.4 Software Manual). Thus, the reproducibility of sample measurement was a much larger factor than the inherent ability of NIR to discern changes among oxidized samples. These results indicate that improved sample presentation techniques which can generate more representative sample spectra can improve the quality of these models.

The miniaturized sample system employed here provided good although not quantitative results for a set of moderately to well oxidized samples (canola set) and poor results for a set of limited to moderately oxidized samples (pecan set). As expected, a combination of the sample sets resulted in a combined model of intermediate quality. Further improvement in sampling techniques to reduce spectral variation and enhance reproducibility could enable NIR to provide quantitative results of conjugated dienes in solid food systems.

5.2.6.2 Peroxide Values

Statistics of the best models for peroxide values are summarized in Table 33. Models shown were based on the full spectral range scanned by the NIR as well as selected wavenumber ranges isolated by the ranking procedure. As evident from the RPD values, limitation of wavenumbers used failed to improve model quality for the pecan oil samples. Although a very limited improvement was observed for canola oil samples when derivatization was used as a pretreatment, significant improvement was observed when only normalization and smoothing were applied. This indicated that the optimization process may be linked to the pretreatment applied. Additional improvement was observed when T0 canola samples were also included. Although full spectrum models for the combined data set were already superior to those of individual data sets, the use of selected wavenumber ranges moderately improved quality.

Table 33. Statistics of best models for peroxide values.

		Calibration		Validation				
Wavenumbers	Q-Value	r^2	SEC	r^2	SECV	BIAS	$SD_{Reference}$	RPD
Pecan Samples								
(Normalization (SNV) and Smoothing)							All Samples Except T0 (n = 18)	
Full Spectrum (4000 - 10,000 cm^{-1})	0.5030	0.5421	18.06	0.4140	21.34	-0.7721	26.64	1.248
Selected Ranges *	0.5058	0.5346	18.21	0.4112	21.41	-0.7736	26.64	1.244
Canola Samples								
(Normalization (SNV), Smoothing and 1st Derivative)							All Samples Except T0 (n = 18)	
Full Spectrum (4000 - 10,000 cm^{-1})	0.4692	0.8514	72.98	0.7289	102.7	3.951	189.3	1.843
Selected Ranges *	0.5142	0.8707	68.08	0.7814	91.91	1.504	189.3	2.060
Canola Samples								
(Normalization (SNV) and Smoothing)							All Samples Except T0 (n = 18)	
Full Spectrum (4000 - 10,000 cm^{-1})	0.4449	0.7739	90.02	0.6109	125.1	5.995	189.3	1.513
Selected Ranges *	0.5968	0.9308	49.82	0.8958	66.03	4.783	189.3	2.867
							All Samples (n = 21)	
Selected Ranges *	0.6511	0.9320	48.88	0.9050	61.15	-0.01690	187.5	3.066
Pecan and Canola Samples								
(Normalization (SNV) and Smoothing)							All Samples Except Canola T0 (n = 39)	
Full Spectrum (4000 - 10,000 cm^{-1})	0.6200	0.9114	45.83	0.8693	58.01	0.09040	154.0	2.654
Selected Ranges *	0.7328	0.9184	43.99	0.8998	50.10	-0.3785	154.0	3.073

* 4176 - 4360 cm^{-1}, 4592 - 4632 cm^{-1}, 5216 - 5260 cm^{-1}, 5768 - 5884 cm^{-1}, and 8808 - 9116 cm^{-1}

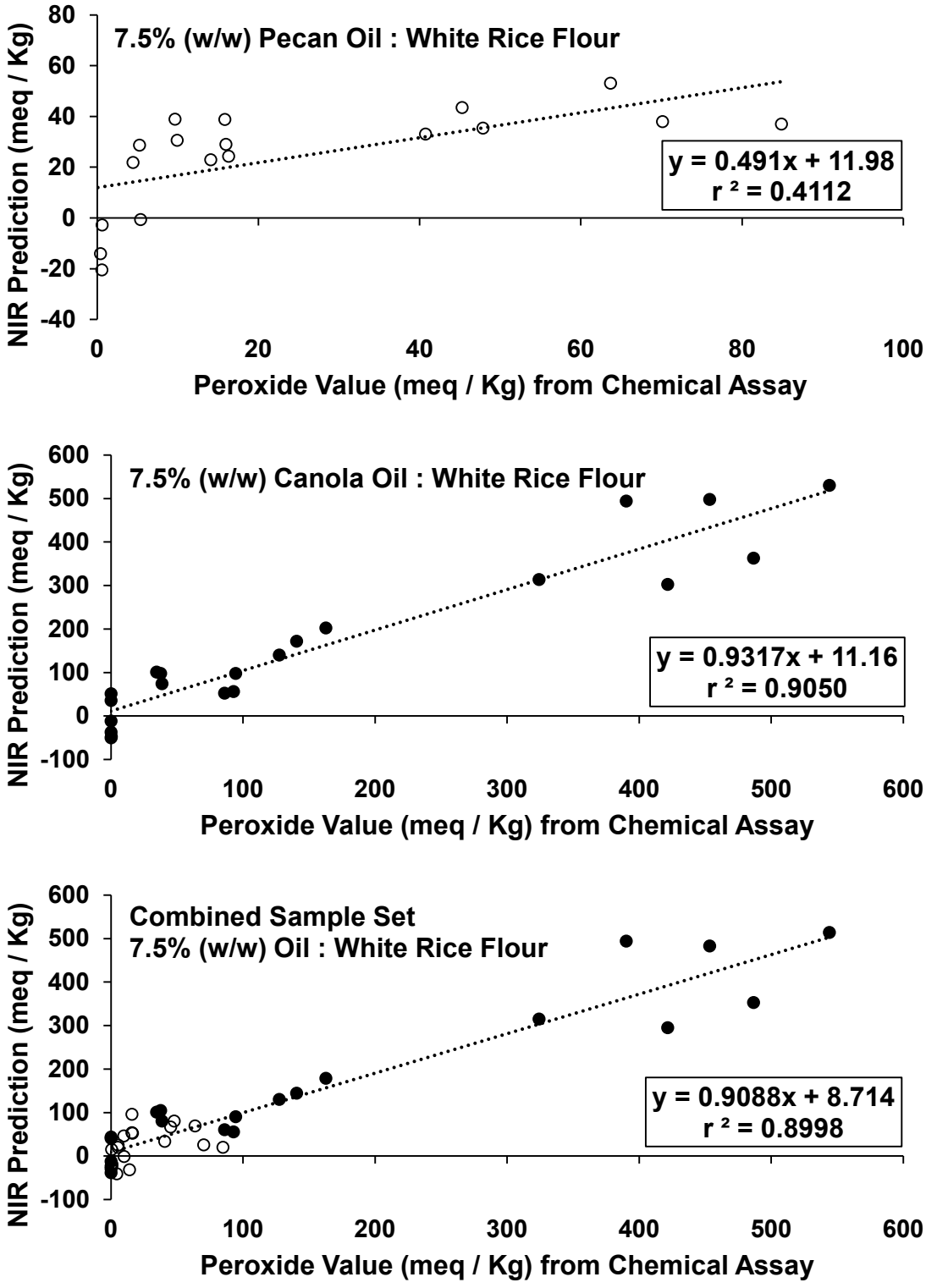


Figure 69. Original vs. NIR predicted peroxide values (meq / Kg) for best models of pecan oil samples (n = 18) (top), canola oil samples (n = 21) (center) and combined sample set (n = 39) (bottom): Pecan oil samples (○); Canola oil samples (●).

Although RPD values of the best models for canola (3.066) and combined sample (3.073) data sets fell between the 'poor' and 'fair' ranges (Williams, 2001), they were significantly larger than the best RPD reported for a peroxide value model in a solid food system (1.42 for moist Asian noodles, Kaddour et al., 2006).

Plots comparing original and NIR predicted peroxide values for the best models for each sample set are shown in Figure 69. Poor linearity was observed among samples made with pecan oil (Figure 69, top), while moderately good linearity was observed for canola oil samples (Figure 69, center) and to a lesser extent for the combined samples (Figure 69, bottom). As was the case with models of conjugated dienes, the disparity in fit can be explained by the degree of oxidation of the respective samples.

Unlike the oxidation index values observed for the two oils during the ASL, a nearly six and a half fold disparity in peroxide values was observed between canola and pecan oils. Given the much greater stability of pecan oil, the inability to obtain a good linear model for pecan samples in this study showed that the error tolerable in the canola models which used more oxidized samples was not acceptable for less oxidized samples with lower peroxide values. Moreover, although models from canola and combined data sets were much better than those of pecan, neither was quantitative indicating error affected these models as well.

Part of this error may have arisen from the reference assay values as the PeroxySafe™ kit (MP Biomedicals, Solon, OH) was subjected to repeated temperature cycling during the course of the ASL, gradually degrading the response of the standard curve. The largest effect would have been on later samples, and unlike those of canola the longest incubated pecan oil samples remained limited in their extent of oxidation. The fact that the peroxide value model was worse than the conjugated diene model for pecan oil samples supports this since the same spectral data set was used with the respective reference assay data sets to generate each model.

It was also likely that the number of replicates used (fifteen tablets scanned three time each) for NIR with the miniaturized sample presentation platform was insufficient to result in an

average spectrum for each sample batch that accurately reflected the true mean spectrum of that batch. As observed in the conjugated diene data, values of SECV were much greater than those of the BIAS across all models (Table 33), reinforcing the finding that reproducibility of sampling was a greater issue than the inherent accuracy of NIR in quantifying lipid oxidation. Thus, in addition to ensuring accurate reference assay values throughout the ASL, further improvements in sample presentation methods may enable NIR to quantitatively model peroxide values as well as conjugated dienes in solid food systems.

5.2.6.3 Interpretation of Wavenumbers

Figure 70 shows the position of important bands in the NIR region pertaining to analytes of interest and interfering agents in lipid oxidation assays. The data are from the Colthup chart in Appendix A. Bands corresponding to CH bonds are widespread and can be linked to conjugated dienes as well as a host of matrix components such as starch from the rice flour used in the samples. Bands corresponding to peroxides have been found in more limited regions, although interpretation can suffer the effects of interference from water and other oxygenated products.

For accurate quantitation of unknowns, NIR models must primarily include wavenumbers pertinent to the analyte of interest. Additionally, wavenumbers corresponding to agents interfering with readings at the primary wavenumbers must be included to overcome the selectivity problem. Wavenumber ranges used in the best quality models for lipid oxidation parameters determined from this study are shown in Figure 71. For each parameter, these models were based upon either the canola sample set or the combined set of canola samples and pecan samples. The pecan sample set alone was insufficient to provide models of sufficient quality for either parameter due to the limited extent of oxidation in the samples assayed.

The best models for conjugated dienes used eight ranges from 4000 - 9100 cm^{-1}. The best models for peroxide values used five ranges which were more narrowly tailored over a slightly smaller span from 4176 - 9116 cm^{-1}. This result generally agreed with the relative prevalence of bands for CH and peroxides (Figure 70). Overlap in the regions used by each model was also significant (Figure 71). This was to be expected as each lipid oxidation parameter varied during the ASL. The model for each oxidation product would be expected to include an accounting for the other as part of a shifting spectral background.

5.2.6.3.1 Models of Conjugated Dienes

Spectral regions used in the best model from this study (Figure 71) aligned well with that expected from the data for CH bands (Figure 70). The regression coefficients for the best models

derived from a set of canola oil samples or a combination set including both pecan and canola oil samples are shown in Figure 72, with wavenumbers of regression coefficient maxima labeled for the canola model.

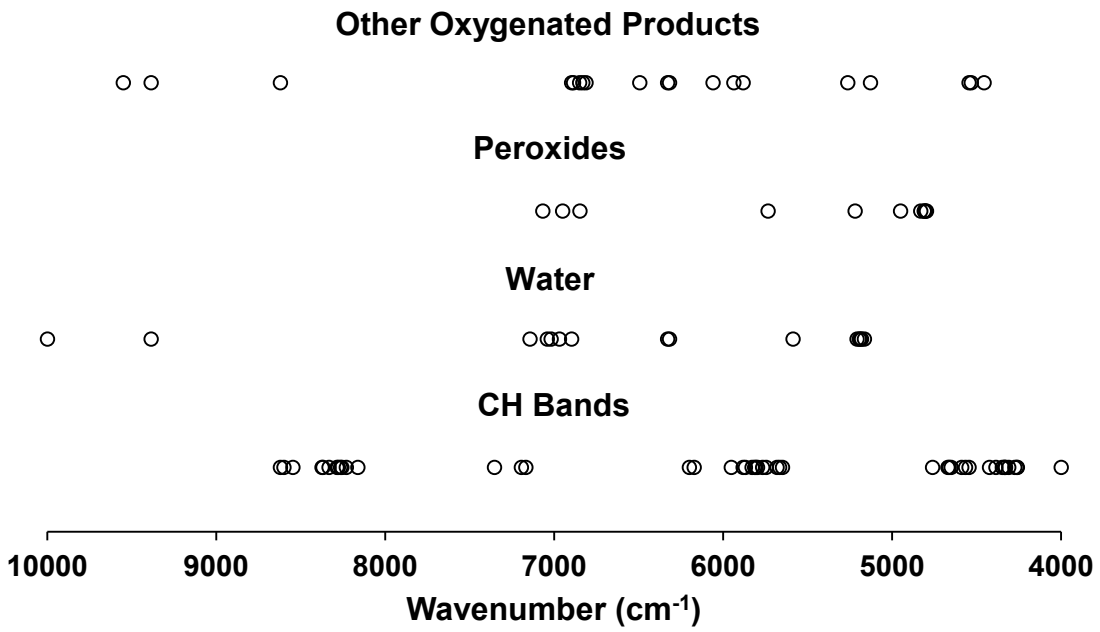


Figure 70. Characteristic wavenumbers important in lipid oxidation assays. Each circle represents a wavenumber corresponding to the group or product specified reported in the literature (see Appendix A).

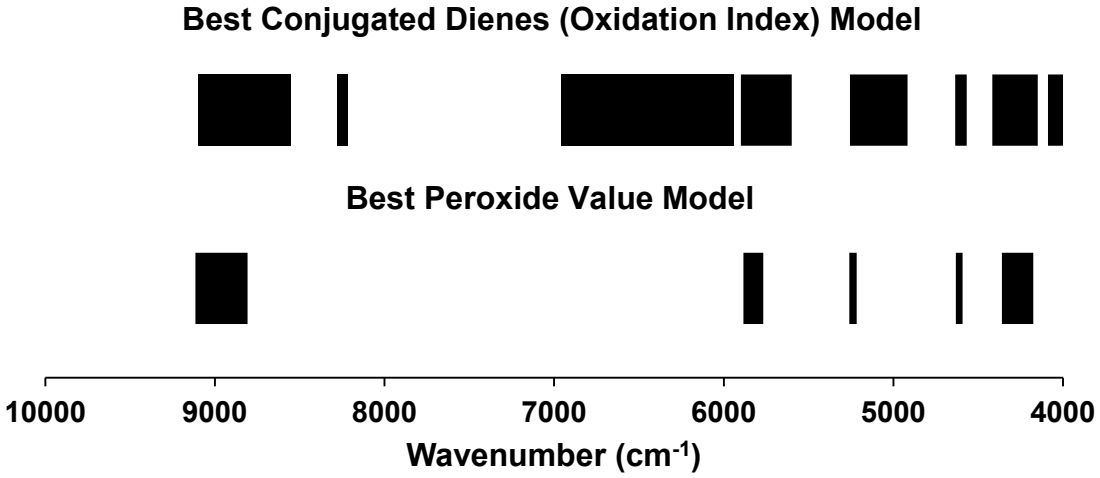


Figure 71. Wavenumber ranges used in the best models for lipid oxidation parameters.

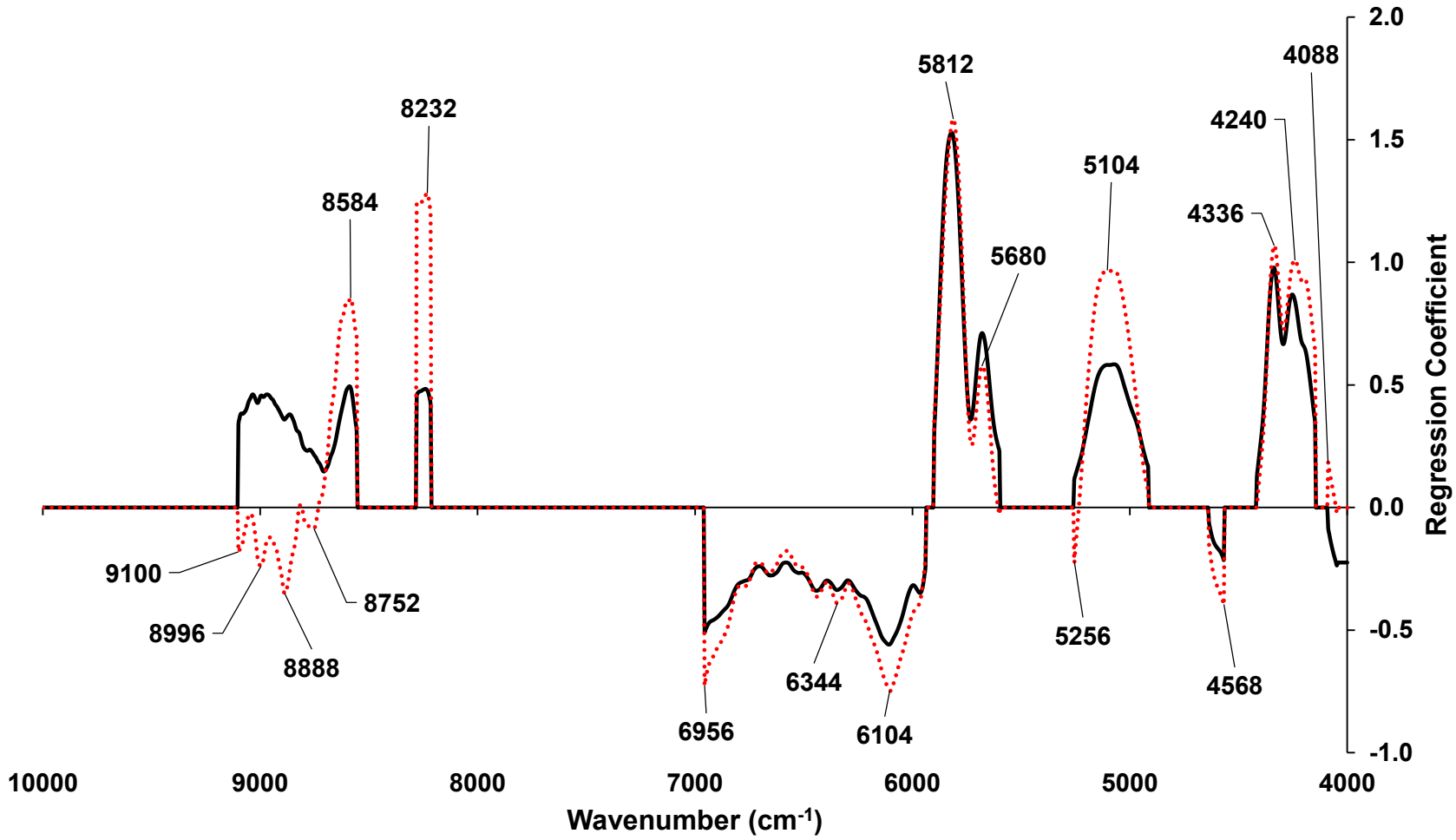


Figure 72. Regression coefficients of best conjugated diene models from 7.5% oil: white rice flour samples: (···) canola samples; (–) combined set of pecan and canola samples.

Assignment of the functional groups important to the conjugated diene models was made by comparison of the Colthup chart data in Appendix A with these regression coefficients (Figure 73). The largest regression coefficients were in the 5600 - 5900 cm^{-1} region around 5812 cm^{-1} (canola model) or 5820 cm-1 (combined pecan and canola model). CH bands are well documented in this region in oxidation studies of numerous oils as shown in Table 34 (reproduced from Appendix A). The experiment to discriminate among lipids indicated this region was significant in distinguishing saturated from unsaturated fats (Figure 22). Figure 73 also indicates methyl and methylene signals in this area. Although carboxylic acid, ketone, alcohol, amine and amide signals also appear in this region, all but the final two of these arise from CH signals associated with these compounds (Appendix A). Thus, as expected, conjugated diene models accord weight to regions indicative of carbon-carbon double bonds.

Table 34. CH bands reported near key wavelengths in best conjugated diene model.

Wavenumber (cm^{-1})	Feature	Compound / Lipid System	Reference
5780	Oil	Ground Beef from Grass or Grain Fed Cattle	Realini et al., 2004
5787	Methylene CH (Asymmetric)	Aliphatic Compounds	Workman, J., & Weyer L. (2008) [Appendix 4a]
5797	C-H Str. first overtone	CH2	NIRCal 5.4
5797	Methylene CH	Aliphatic Compounds	Workman, J., & Weyer L. (2008) [Appendix 4a]
5797	CH2 (more intense for fully saturated fatty acids)	Palm; Rapeseed	Kaddour et al., 2006
5807	CH	Soybean Oil	Cho et al., 1998
5807	CH2	Marine & Linseed Oil Dietary Supplements	Berzaghi et al., 2005
5814	CH	Fish Oil	Cozzolino et al., 2005
5829	CH	Olive; Maize; Seed; Sunflower	Armenta et al., 2007

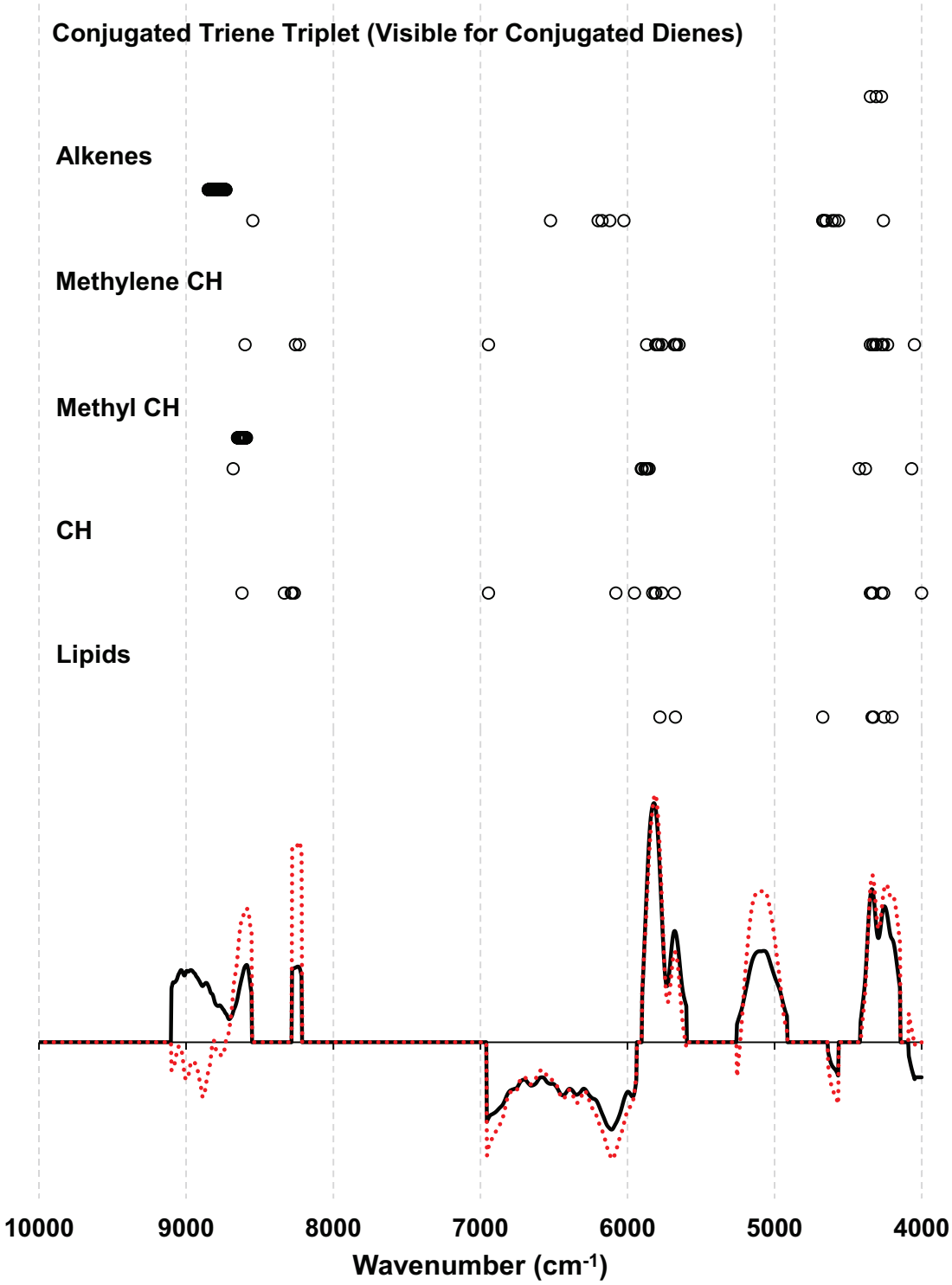


Figure 73. Assignment of functional groups to important wavenumbers of conjugated diene models. Each circle represents a wavenumber corresponding to the group or product specified reported in the literature (see Appendix A).

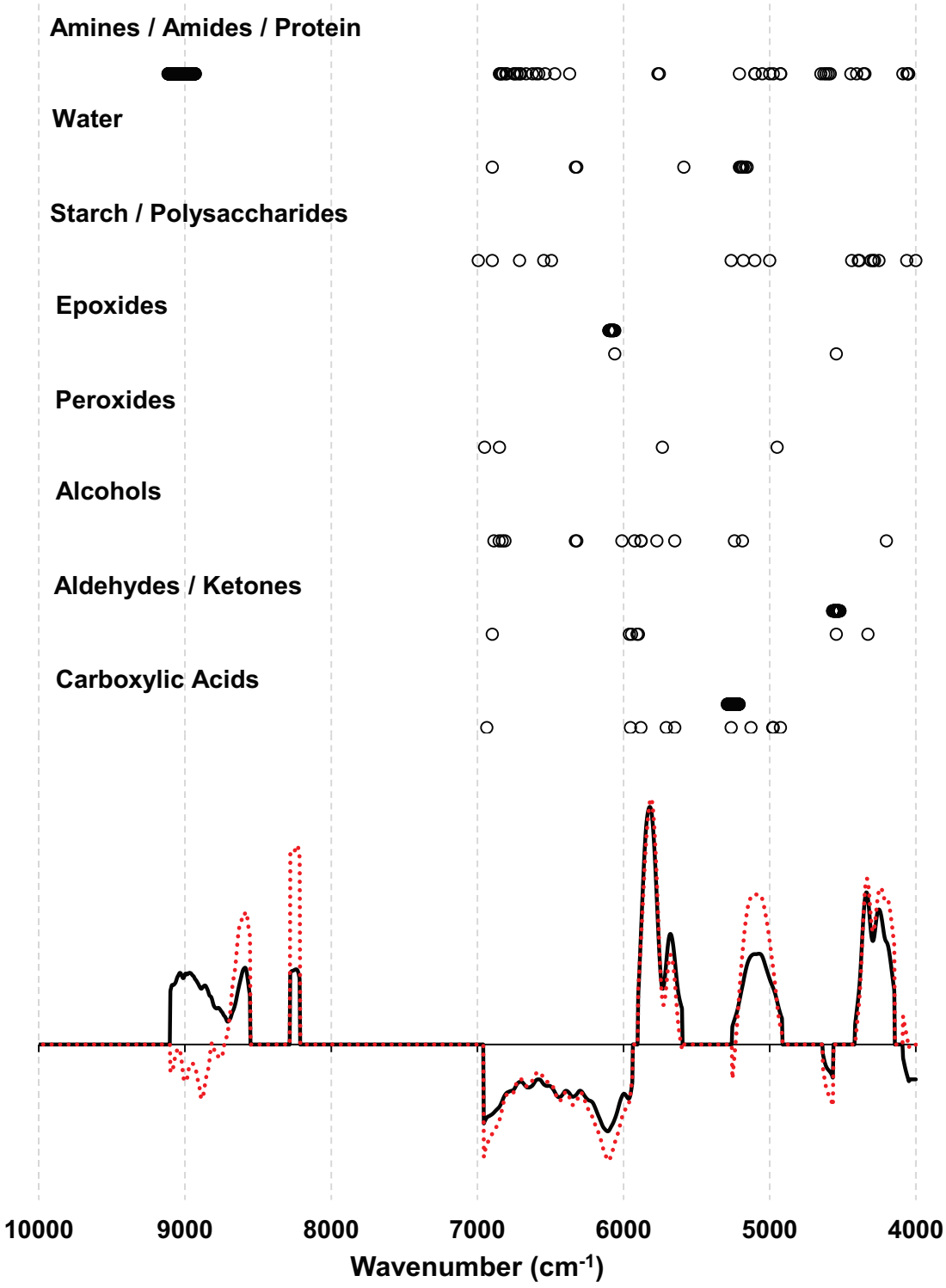


Figure 73. (Continued) Assignment of functional groups to important wavenumbers of conjugated diene models. Each circle represents a wavenumber corresponding to the group or product specified reported in the literature (see Appendix A).

Another significant range in both conjugated diene models was 4148 - 4416 cm^{-1}. This region was linked to lipids and CH bands in numerous oxidation studies (Figure 73 and Appendix A) as well as an alkene signal at 4261 cm^{-1} (NIRCal 5.4 Software). The conjugated triene triplet observed in standards by Holman and Edmondson (Holman & Edmondson, 1956), also evident in trans, trans conjugated dienes, was associated with this range as well. A number of signals corresponding to starch and amide bonds have also been found in this area. This region was likely important for determining the relative amount of oil to flour for any given sample because signals for lipid, carbohydrate and protein constituents have each been found here. The determination that much of this region was also included in the best peroxide value models (see below) reinforced this role. Thus, 4148 - 4416 cm^{-1} was used for direct determination of conjugated dienes as well as adjustment of the model for sample composition.

The next significant range included a single broad peak from 4916 - 5256 cm^{-1}. Regression coefficients in this range differed among models, with those of the canola sample model being much larger than the model using both pecan and canola samples (Figure 72). No hydrocarbon CH or alkene signals were found that overlap this region, but numerous oxygenated products including carboxylic acids, alcohols, polysaccharides and water as well as proteins have been determined here (Figure 73 and Appendix A). The secondary wavelength (4950 cm^{-1}) in a MLR model of peroxide values in soybean oil (Cho et al., 1998) also fell within this range. This region was most likely the key indicator of moisture content among samples, although a second role is needed to explain the discrepancy in regression coefficients among models with and without pecan samples. Shifting amounts of other oxygenated products in less oxidized (pecan) and more oxidized (canola samples) may have caused this difference, indicating a dual purpose for this range. Thus, to ensure accuracy in lipid oxidation measurements the moisture content of samples should be monitored as well.

Other significant wavenumber ranges where regression coefficients differed among models were 8216 - 8280 cm^{-1} and 8556 - 9100 cm^{-1}. Only methylene CH signals have been

found in the relatively narrow former region (Figure 73 and Appendix A), thus it is another indicator of the degree of unsaturation. Changing oxidation levels among the two oils as well as differences in fatty acid composition [oleic versus linoleic and linolenic acid content (see Table 17)] could explain the discrepancy in regression coefficients among models. The initial peak (8584 cm^{-1}) in the latter region was also clearly associated with oxidation, as corresponding methyl and methylene CH and alkene bands (Figure 73 and Appendix A) indicated the degree of unsaturation. An aliphatic carbonyl signal has also been reported nearby at 8621 cm^{-1} indicating the possible contribution of aldehydes or ketones (Appendix A). The broad set of additional peaks (8700 - 9100 cm^{-1}) overlapped regions documented for the second overtone of alkenes as well as hydrogen bonded secondary amides indicative of proteins (Figure 73 and Appendix A). The highest energy (shortest wavelength or longest wavenumber) area included in the models, this region would have been most affected by application of the Fourier transformation as well as differences in pressure among the samples (see pressure discrimination model loadings in Figure 30).

The model also included wavenumbers in the 5940 - 6956 cm^{-1} range with two peaks flanking the region at 6104 cm^{-1} and 6956 cm^{-1}. Multiple signals have been found in this area for matrix constituents (starch, water and proteins) as well as lipid oxidation products including alkenes, carboxylic acids, alcohols, peroxides and epoxides (Figure 73 and Appendix A). Functional groups with signals closest to the major peaks in the region include CH (6079 cm^{-1}), alkenes (6120 cm^{-1}) and epoxides (up to 6100 cm^{-1}) on the low energy side and methylene (6944 cm^{-1}), carboxylic acids (6935 cm^{-1}), peroxides (6950 cm^{-1}), starch (6993 cm^{-1}) and water (6897 cm^{-1}) on the high energy side.

Finally, two minor regions were also included in models of conjugated dienes. The peak (4568 cm^{-1}) of the narrow range from 4568 - 4636 cm^{-1} coincided with cis unsaturation signals (4566 cm^{-1}) found in standards (Holman & Edmondson, 1956; Holman et al., 1958) and breast meat from chickens fed supplements of marine and linseed oils (Berzaghi et al., 2005). An

alkene combination band (Wojcicki et al., 2015), proteins and aldehydes have also been found in this range (Figure 73 and Appendix A), as well as a nearby band corresponding to terminal epoxides (4545 cm^{-1}) (Goddu & Delker, 1958). The second minor region included 4000 - 4088 cm^{-1}, corresponding to starch and protein signals as well as methylene groups.

5.2.6.3.2 Models of Peroxide Values

The best models of peroxide values relied upon many fewer wavenumbers in five narrow ranges nearly entirely overlapped by ranges used in models of conjugated dienes (Figure 71). The regression coefficients for the best models derived from a set of canola oil samples or a combination set including both pecan and canola oil samples are shown in Figure 74 with wavenumbers of regression coefficient maxima labeled for the canola model.

Assignment of the functional groups important to the peroxide value models was made by comparison of the Colthup chart data from Appendix A with these regression coefficients (Figure 75). The largest regression coefficients were confined to two regions, 5216 - 5260 cm^{-1} and 5768 - 5884 cm^{-1}. The former overlapped the high energy side of a broad region used in conjugated diene models. The peak value at 5256 cm^{-1} corresponded to signals for carboxylic acids (5263 cm^{-1}), alcohols (5241 cm^{-1}) and starch (5263 cm^{-1}) (Figure 75 and Appendix A). Also in this region, in NIR studies of peroxides in canola, olive and safflower oils, Takamura (Takamura et al., 1995) ascribed a signal at 5219 cm^{-1} to a possible secondary oxidation product.

The region from 5768 - 5884 cm^{-1} was identical to the area with the highest regression coefficients among conjugated diene models. CH signals have been found throughout this area, while methyl groups in alcohols have been found near the low (5773 cm^{-1}) and high ends (5880 - 5882 cm^{-1}) and those in acids have also been found (5882 cm^{-1}) (Figure 75 and Appendix A). Although Takamura (Takamura et al., 1995) determined 4798 cm^{-1}, well outside this region, to be the most important wavenumber for monitoring peroxides via NIR, correlation with peroxide values was also found at 5734 cm^{-1}, also outside but near the low end of this region.

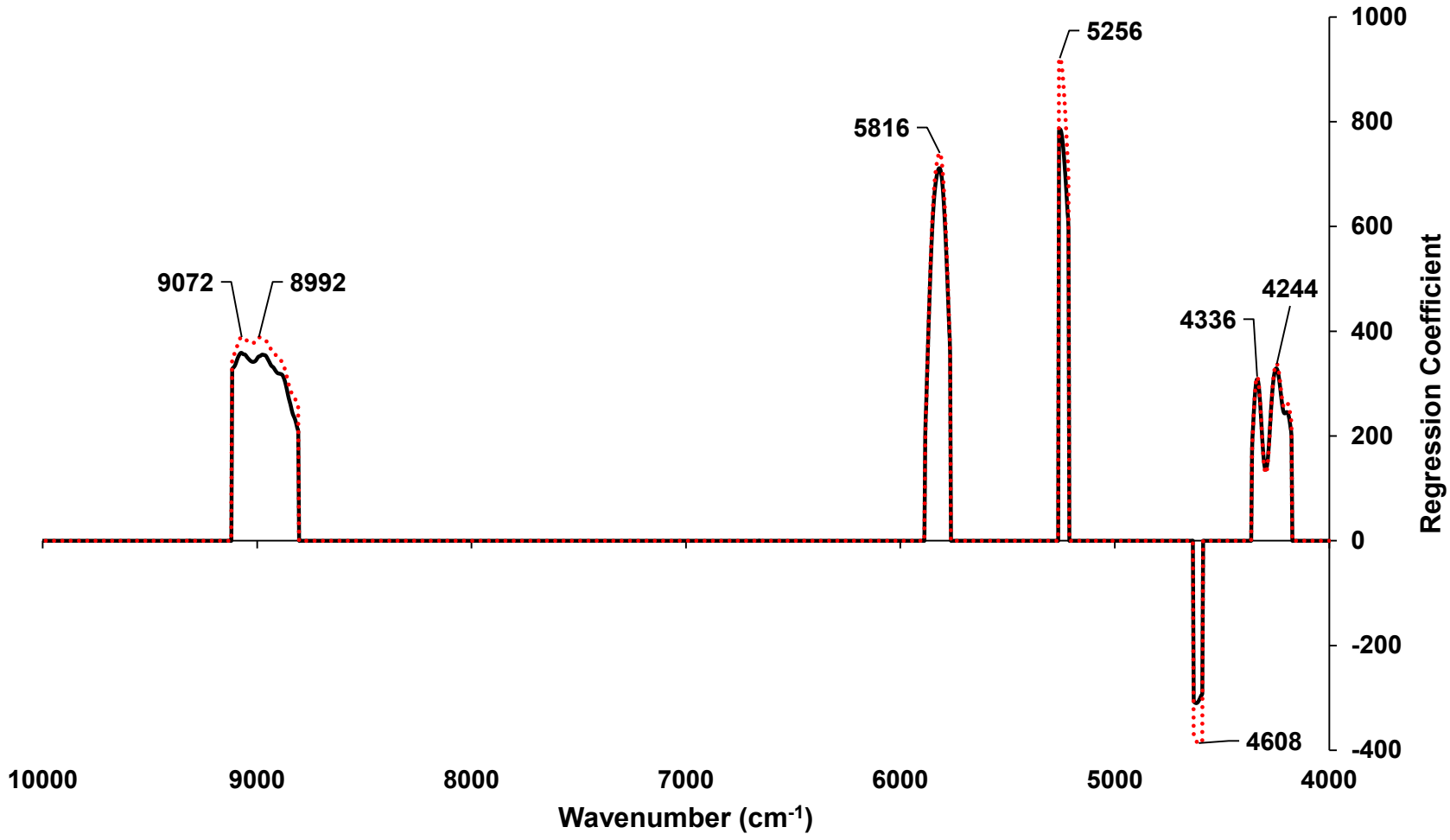


Figure 74. Regression coefficients of best peroxide value models from 7.5% oil: white rice flour samples: (···) canola samples; (–) combined set of pecan and canola samples.

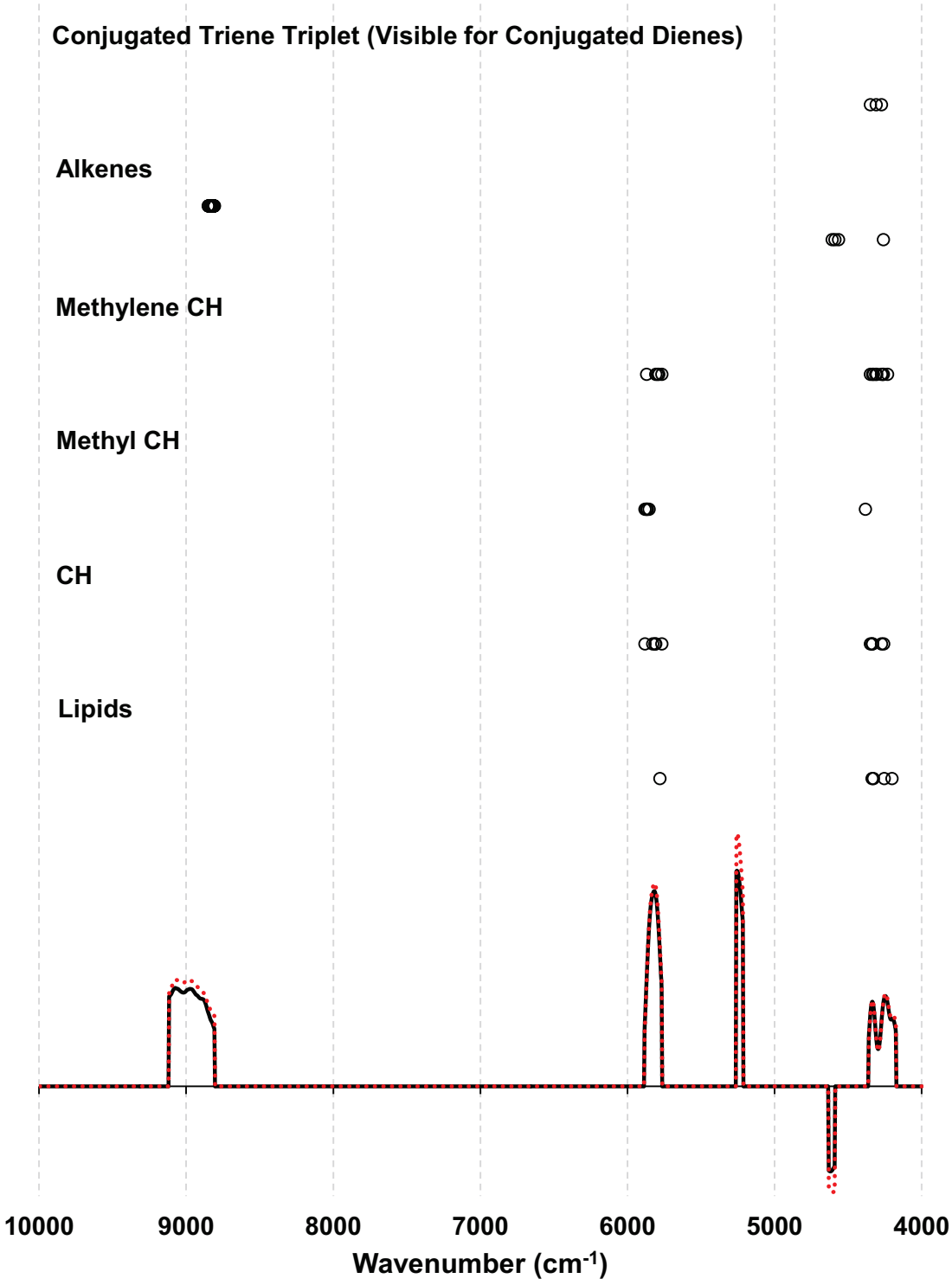


Figure 75. Assignment of functional groups to important wavenumbers of peroxide value models. Each circle represents a wavenumber corresponding to the group or product specified reported in the literature (see Appendix A).

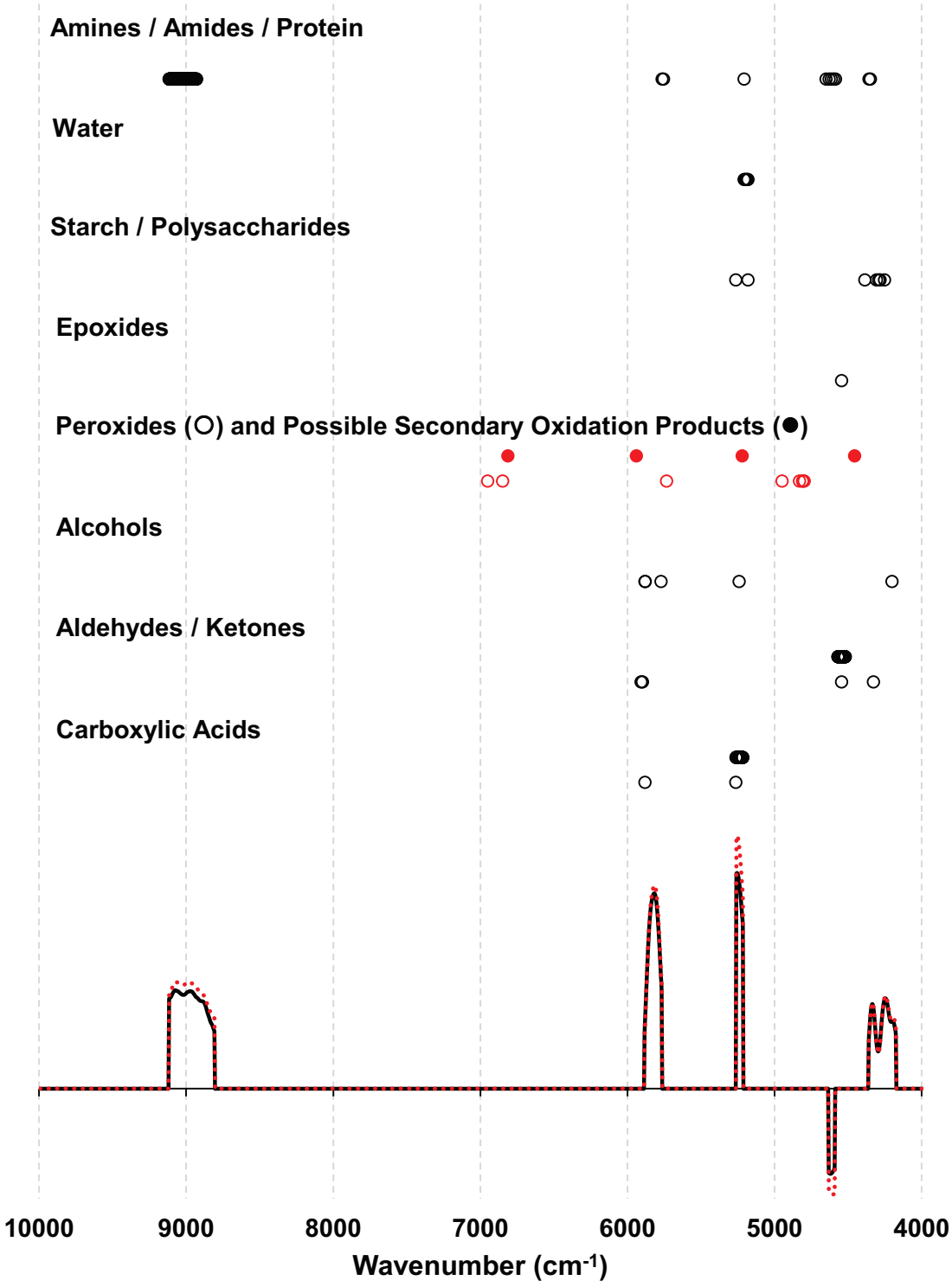


Figure 75. (Continued) Assignment of functional groups to important wavenumbers of peroxide value models. Each circle represents a wavenumber corresponding to the group or product specified reported in the literature (see Appendix A).

Regression coefficients in the remaining three wavenumber ranges were significantly reduced. Two of these ranges, 4176 - 4360 cm^{-1} and 8808 - 9116 cm^{-1}, were slightly truncated but had similar peak values relative to corresponding regions in the conjugated diene models. Canola model regression coefficients in the latter range had opposite signs for conjugate diene (Figure 72) and peroxide value (Figure 74) models. However, T0 and final week samples were excluded from conjugated diene models but not those of peroxide values. Thus, differences in wavenumber ranges as well as samples used could explain the shift. As neither wavenumber range was associated with peroxide signals, each was likely incorporated to monitor matrix composition and other background changes (fluctuation of alkene and conjugated diene signals) arising from lipid oxidation according to the functional group assignments made in the discussion on conjugated diene models (see Section 5.2.6.3.1).

The final range, 4592 - 4632 cm^{-1}, was also slightly truncated from a corresponding region (4568 - 4636 cm^{-1}) in the conjugated diene models, although the peak value in the peroxide models shifted to 4608 to 4620 cm^{-1}. Despite loss of the band for cis unsaturation from the conjugated diene model, bands for alkenes and proteins (hydrogen bonded secondary amides) remain in this narrower region (Figure 75 and Appendix A). The range also included bands reported in an assay of fried sunflower oil for linoleic acid content (4604 cm^{-1} and 4621 cm^{-1})(El-Rafey et al., 1988), a significant fatty acid in both canola and pecan oils (Kostik et al., 2013; Toro-Vazquez et al., 1999). Although aldehydes and terminal epoxides have also been reported nearby at 4545 cm^{-1}, the nearest peroxide signal reported fell well above the region at 4798 cm^{-1}.

Interestingly, none of the wavenumbers from the 4800 - 5000 cm^{-1} range in which peroxides have been reported in lipid oxidation studies using NIR [4798 cm^{-1} (Takamura et al., 1995); 4808 cm^{-1} and 4950 cm^{-1} (Cho et al., 1998); 4810 cm^{-1} (Wojcicki et al., 2015); and 4831 cm^{-1} (Holman & Edmondson, 1956; Holman et al., 1958; Yildiz et al., 2001)] were isolated in the best peroxide value models in this study. Also, wavenumbers near the center of the NIR spectrum were fully dismissed despite reports of peroxides at 6849 cm^{-1} (Holman & Edmondson,

1956; Holman et al., 1958) and 7068 cm^{-1} (Wojcicki et al., 2015). These were likely due to the manner in which the ASL study was conducted and the sample presentation apparatus used.

Unlike prior studies of lipid oxidation using NIR spectroscopy, in this investigation neat oils were oxidized and then mixed with a solid carrier to a nominal composition of 7.5% w/w oil prior to the acquisition of NIR spectra. This procedure ensured that correlations would arise from lipid oxidation in the oil per se rather than interactions with or oxidation of components of the rice flour matrix during thermal treatment. It also enabled the reference assay to be performed without the complicating effects of extraction from the solid matrix. However, to perform NIR spectroscopy on a solid sample, the oxidized oil first needed to be mixed with the rice flour matrix.

NIR sample preparation was accomplished using food grade (Type 316) stainless steel mixing equipment, which includes iron, chromium, manganese and molybdenum ("Stainless Steel 304-Alloy Composition", 2019). Although in stainless steel these metals are present in their elemental state, ionized forms are possible via corrosion. Isopropyl alcohol was used to remove lipids when this equipment was cleaned between samples. The compatibility between this solvent and stainless steel (types 304 or 316) is rated as "B-Good: Minor Effect, slight corrosion, or discoloration" ("Chemical Compatibility Database", 2019). In the presence of reducing metals such as Fe^{2+}, lipid hydroperoxides (LOOH) can degrade to alkoxyl radicals (LO•) and hydroxide ions (Schaich, 2005).

It is also possible for lipid alkoxy and peroxyl radicals to undergo co-oxidation with proteins and carbohydrates (Schaich, 2008). Both of these effects could have occurred to account for the loss of a clear hydroperoxide signal in the solid sample despite the high degree of correlation among NIR and reference assay results in models using canola oil samples where high levels of lipid hydroperoxides were attained during the ASL. Lipid hydroperoxides in neat oils had the opportunity to decompose to other oxygenated products, such as alcohols, aldehydes and carboxylic acids, when those oils were incorporated into white rice flour.

Regarding the absence of signals near the middle of the NIR where peroxides have been reported, in addition to the considerations above it is known that moisture affects this spectral region (Appendix A). Fluctuations in moisture content can increase variability, decreasing model quality and hindering variable selection in improvement techniques. High variability among pecan and canola samples of the combined data sets was observed at and above 7000 cm^{-1} in both conjugated diene and peroxide value models (see bottom of Figures 56 and 57).

The certificate of analysis for the rice flour used in all samples indicated it contained 9.2% moisture. Rice flour is hygroscopic but was stored under argon in Ball jars sealed with gas impermeable Teflon tape. Although exact moisture control was not possible during sample preparation and analysis, these were carried out in a small (roughly 50 sq. ft.) dedicated laboratory on the top floor of the building. A review of campus meteorological records for days on which apparent outliers were assayed indicated no unusual weather or major shifts in relative humidity. Thus, while moisture content is important to control, in this study an alternative explanation was needed for the observed variability.

Boron oxide is the agent responsible for limiting thermal expansion in borosilicate glass. Hanst et al. (1965) observed a broad infrared peak for boron oxide vapor centered near 1400 cm^{-1}. Multiples of this fundamental include 7000 cm^{-1}. The specifications for the borosilicate glass discs used here indicated a thickness of 2 +/- 0.2 mm and a diameter of 1.125” +/- .01”. For each sample tablet, discs were seated in the insert for placement in the rotating sample holder. Discs may not have been perfectly flat depending on the fit within the holder insert. Moreover, samples were friable and small pieces of flour were able to fall between the lip at the base of the insert and the bottom of the disc. These problems were exacerbated because the assay used transflectance measurements. Radiation needed to travel twice through the disc, once to reach the sample and a second time to return to the detector. These changes in placement of the discs may have added significant variability by altering the path length for radiation to travel through the borosilicate glass substrate on a sample by sample basis.

The effect of the borosilicate glass substrate on pretreated (normalized and smoothed) difference spectra for the fifteen original sample tablets of pecan week 1 sample batch A is shown in Figure 76. An average spectrum for the batch was constructed from the fifteen sample tablets and subtracted from the average spectrum for each tablet to generate the difference spectra shown. The spectrum highlighted in red was for a tablet erroneously inserted in the holder without a borosilicate glass disc. Clearly, the region just above 7000 cm^{-1} is among the most greatly affected by incorporation of a borosilicate glass substrate during sample presentation. Based upon the size of the peak in the highlighted spectrum relative to the variability among the remaining spectra in which glass discs were used, variation in placement of the borosilicate glass discs for each sample were the likely reason that wavenumbers in that region were discarded. Accordingly, models could be improved by the use of quartz substrates or a different sample presentation platform (i.e. top down rather than bottom up) not requiring radiation to pass through a substrate.

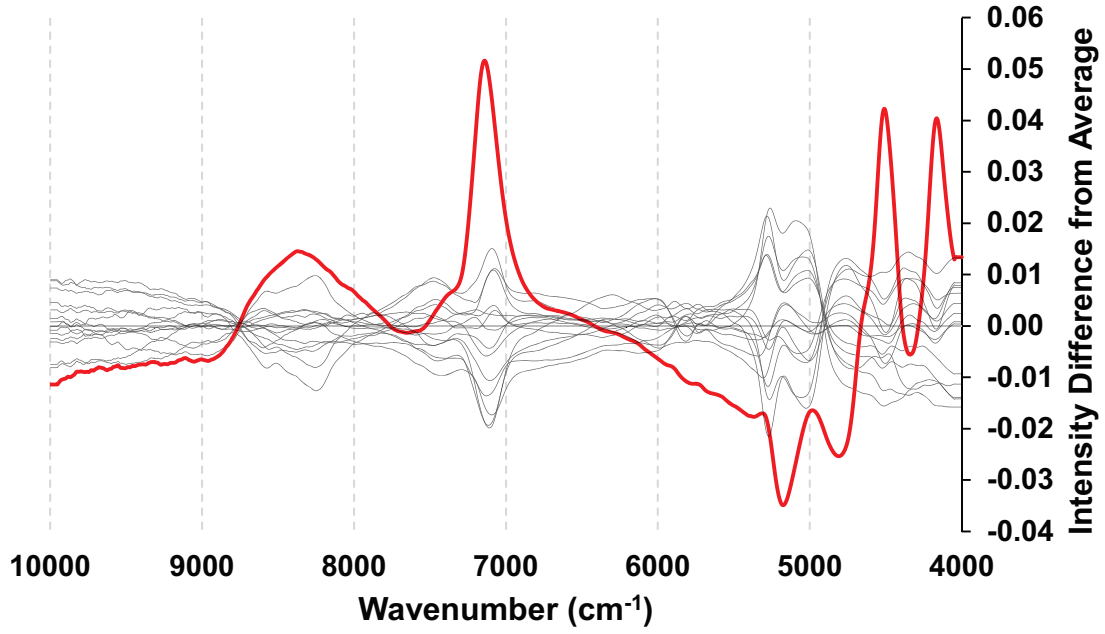


Figure 76. Effect of borosilicate glass substrate on pretreated difference spectra of pecan samples scanned with (-) and without (-) a glass substrate in the path of the NIR beam.

5.2.6.4 Summary of Tentative Functional Group Assignments in Models

Wavenumbers with high regression coefficients in the best models obtained herein correlated with a number of documented signals for functional groups and compounds implicated in lipid oxidation per se as well as constituents of the white rice flour matrix. A summary of the tentative functional group assignments for NIR models of conjugated dienes is presented in Figure 77. A summary of the tentative functional group assignments for NIR models of peroxide values is presented in Figure 78. Sources for each functional group assignment are listed in Appendix A.

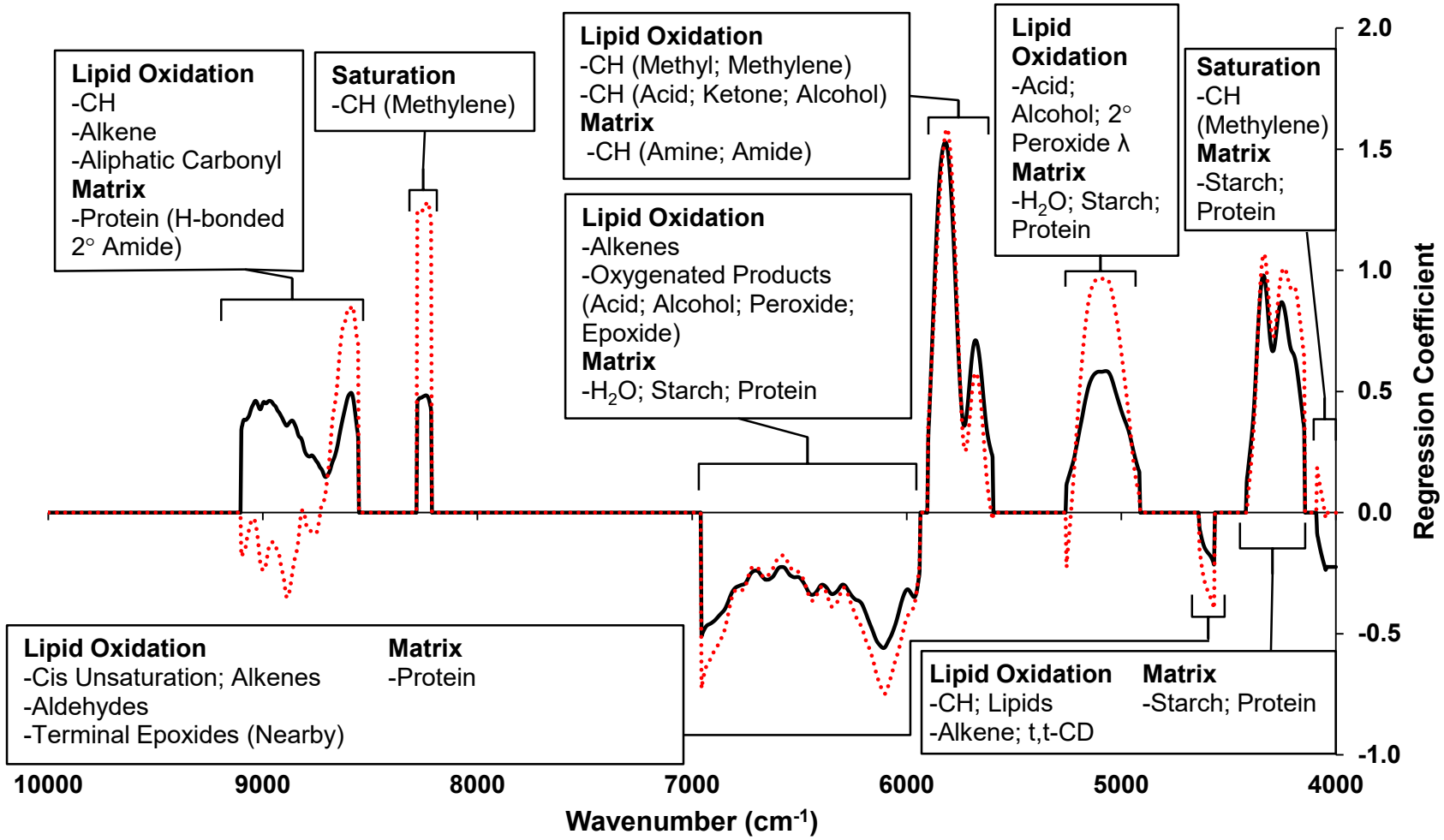


Figure 77. Summary of tentative functional group assignments for important wavenumbers from the best models for conjugated dienes from 7.5% oil: white rice flour samples: (···) canola samples; (–) combined set of pecan and canola samples. Sources of assignments are presented in Appendix A.

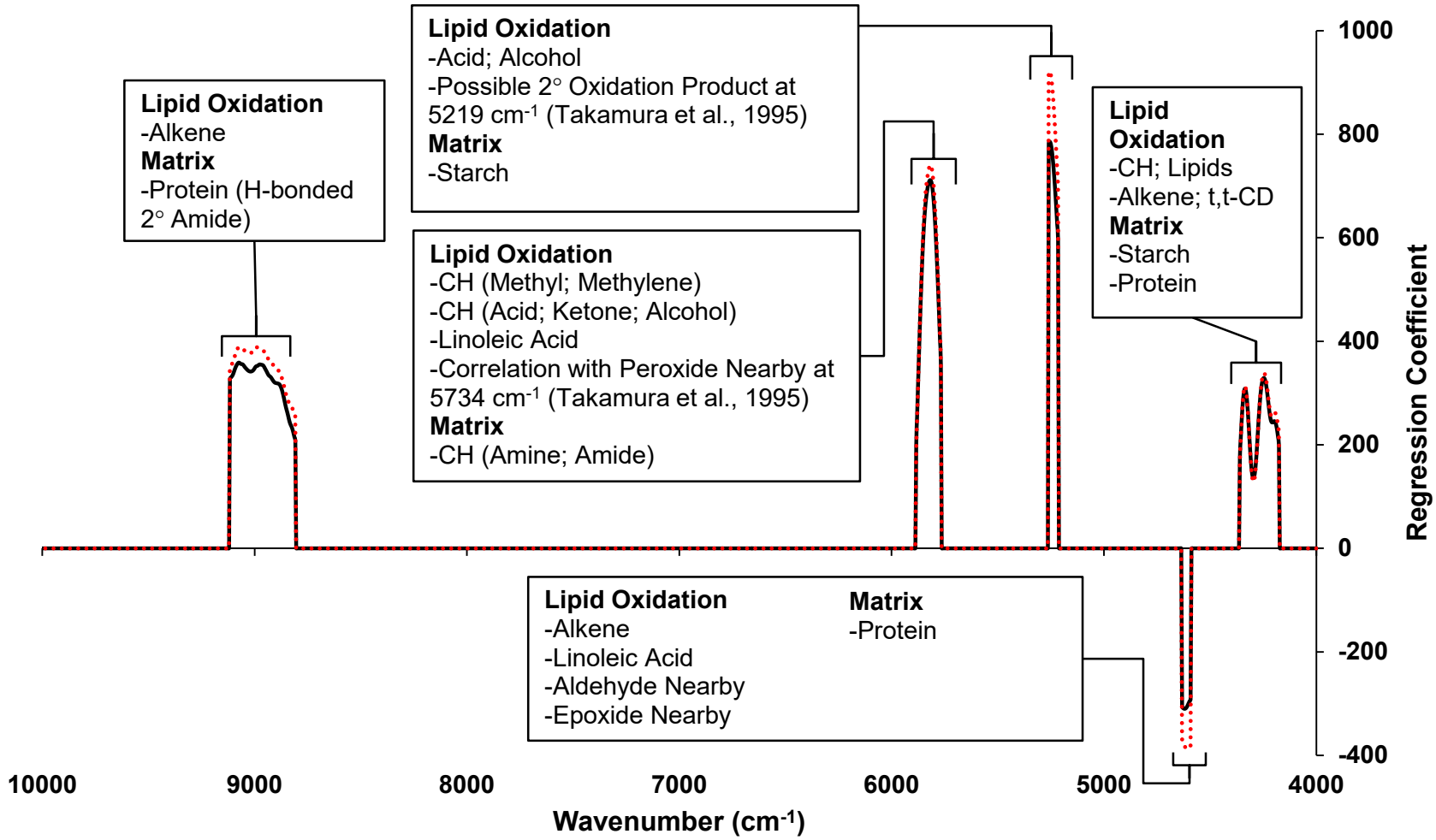


Figure 78. Summary of tentative functional group assignments for important wavenumbers from the best models for peroxide values from 7.5% oil: white rice flour samples: (···) canola samples; (–) combined set of pecan and canola samples. Sources of assignments are presented in Appendix A.

6. SUMMARY AND CONCLUSION

Two main objectives were addressed in this project to determine whether NIR could be applied successfully in studies of lipid oxidation in solid foods. The first was to determine whether the poor performance reported for such assays was due to inherent limitations of the technique or arose from causes that could be remedied to enable application of the technique for rapid analysis of oxidation products. The second was to determine whether it was possible to carry out such assays using sample sizes practical for a research laboratory.

Qualitative assessments were initially made regarding the ability of NIR to discriminate among different fats and oils present as a minor ingredient in a model solid food system and to discriminate among different oxidation levels in oils present as a minor ingredient in a model solid food system. The effect of sample preparation and presentation on variation observed in the initial studies was also investigated. Finally, the ability of NIR to accurately quantify three markers of lipid oxidation, conjugated dienes, lipid hydroperoxides and soluble carbonyls, in oxidized oils present as a minor ingredient in a model solid food system was studied.

NIR was able to discriminate among seven of eight fats and oils assayed at 15% (w/w) in white rice flour. Samples which incorporated chicken fat, lard, or palm, flaxseed, canola or soybean oil were each distinguishable. Sunflower or safflower oil were also distinguishable from the first six but not each other due to similarities in their fatty acid compositions. The model to discriminate all but safflower oil samples emphasized the wavenumber range from 5770 - 5900 cm^{-1}, corresponding to CH stretching vibrations for methyl and methylene groups and indicative of the degree of unsaturation of constituent fatty acids. A scores plot indicated variation along an axis corresponding to the fat or oil content within each sample, as well as random scatter among samples of saturated fats from inconsistencies in sample surfaces due to adhesion of those samples to the die in which they were pressed.

NIR was also able to discriminate among different oxidation levels of the same oil assayed at 15% (w/w) in white rice flour provided samples were assayed at times which

correlated with the duration over which the oil oxidized. Following oxidation of thin films of oil at 60°C for one, two or three weeks, all four time points (including fresh oil at T0) were distinguishable in samples incorporating canola oil, while only the first three time points were distinguishable in samples incorporating soybean oil and only the first two were distinguishable in samples incorporating either sunflower oil or safflower oil. This reflected the different time required for each oil to degrade, as soybean, sunflower and safflower oils had high levels of polyunsaturated fatty acids (linoleic acid (18:2) content >50% in soybean oil and >67% in sunflower and safflower oils) and canola oil was the most stable. Thus, a sampling plan to develop NIR models of lipid oxidation must account for the rate of oxidative degradation of the fat or oil of interest.

Variation among nominally identical samples was evident in both discrimination experiments. Given its negative impact on quantitative assays, sample preparation and presentation methods were assessed to minimize variation. Results of the qualitative NIR model for discriminating oxidation states in canola oil samples indicated variation occurred along the same axis at each time point. Scores of replicates of each sample drifted consistently along the first principal component axis, indicating a systematic temperature or pressure effect degraded the quality of the model. Scores of the first replicate of each sample were randomly distributed along both principal component axes, indicating spurious interferences arising from differences in composition and/or pressure applied to form the samples. To examine the effects of these interferences, a lever press capable of applying a uniform pressure to form the samples and a miniaturized rotating sample platform to average compositional effects were constructed.

Using nominally identical samples of 15% (w/w) canola oil on white rice flour, NIR was able to discriminate among forming conditions at moderate pressures (340 psi, 365 psi and 390 psi) but not those high enough (415 psi) to induce pooling of oil within samples. The hydrophilic nature of the rice flour matrix resulted in sample inhomogeneity induced by applying too great a forming pressure for the amount of oil present in a given sample. Reduction of sample oil content

to 7.5% (w/w) indicated NIR was able to discriminate forming conditions at higher pressures as well. Although loadings from the pressure discrimination model indicated wavenumbers in the 9000 - 10,000 cm^{-1} range were most affected, important wavenumbers from the model for discriminating oxidation states of canola oil were also impacted. Thus, pressure differences among identical samples can give rise to NIR spectral variation in lipid oxidation assays and must be minimized.

A tradeoff existed between the use of pressed and powdered samples. Pressing resulted in improved protection against the environment (moisture, oxygen and light) during scanning on the one hand and increased variability from temperature fluctuations and inhomogeneity induced by the forming process on the other. The latter was compensated for by rotation of the sample during spectral acquisition to increase the number of positions analyzed. Due to the labile nature of lipid oxidation products, avoidance of exposure was critical and thus the accelerated shelf life study was conducted using pressed and rotated samples.

Neat oils were oxidized during an ASL study at 40°C for up to 14.14 to 15.86 weeks for pecan and canola oils, respectively. Conditions were chosen to avoid shifts in oxidation pathways altering the distribution of products normally occurring in foods by elevated study temperatures. Although early products of oxidation were present throughout the incubation, soluble carbonyls only developed in the final time points of canola oil. Thus, NIR models were confined to conjugated dienes and lipid hydroperoxides using three sample sets: samples incorporating pecan oil, samples incorporating canola oil and a combined data set including both types.

The best full spectrum NIR model of conjugated dienes in pecan samples was very poor ($Q = 0.6266$), while that in canola samples was better but not quantitative ($Q = 0.7252$) and that of the combined sample set was of intermediate quality ($Q = 0.7214$). The best full spectrum NIR models of peroxide values were much worse than the corresponding models of conjugated dienes ($Q = 0.5030$ for pecan samples, 0.4692 for canola samples, and 0.6200 for all samples).

Four wavenumber selection methods were investigated to attempt to improve the quality of the best models using a combined set of pecan and canola samples. For both conjugated diene and peroxide value models, ranking of the cross validation regression coefficients t test values provided the greatest improvements. Selection of wavenumbers via this approach provided a moderate improvement of the conjugated diene model (Q = 0.7214 rose to 0.7409) and a significant improvement of the peroxide value model (Q = 0.6200 rose to 0.7328). Selection of certain ranges isolated by the process further improved the conjugated diene model (Q = 0.7519).

Despite reliance on identical NIR sample spectra, conjugated diene models improved as more samples were excluded while peroxide value models favored the opposite. The importance of a broader array of wavenumbers to conjugated dienes, due to the prevalence of CH bands throughout the NIR, afforded more opportunities for interference from a variety of sources to affect models for that product. Peroxide value models relied upon fewer wavenumbers, making them more robust to interference. However, inclusion of far more wavenumbers than necessary in full spectrum models diminished their quality overall relative to corresponding models of conjugated dienes. This coincided with previous reports that single wavenumber models provided good results for peroxide values in neat oils.

Wavenumbers isolated from the improvement process were also applied to the best models based upon sample sets of either oil. Among pecan oil samples, negligible improvement was observed in both conjugated diene (Q = 0.6266 rose to 0.6331) and peroxide value (Q = 0.5030 rose to 0.5058) models. For canola oil samples, significant improvement was seen in the conjugated diene model (Q = 0.7252 rose to 0.7721). Although a similar improvement was observed in the best full spectrum peroxide value model of canola samples (Q = 0.4692 rose to 0.5142), an even greater improvement (Q = 0.4449 rose to 0.5968) was seen in a suboptimal full spectrum model which relied upon the same normalization and smoothing pretreatments used for the combined sample model subjected to the improvement procedure. In keeping with the tendency toward sample inclusion, the optimized peroxide value model for canola samples was

further improved (Q = 0.6511) by reinstatement of the T0 samples. These samples could not be reinstated in the conjugated diene model without degrading its quality. The dramatic improvement in the peroxide value model using normalized and smoothed spectra of canola samples indicated the pretreatment applied affected the outcome of the improvement process.

The sample set used in the model subjected to improvement also affected the wavenumbers selected by the process. Direct improvement of the best model of peroxide values from thirty-nine pecan samples and canola samples resulted in the loss of all wavenumbers from 4500-5000 cm^{-1}, a region known to be associated with peroxides. Improvement of a lower quality model for peroxides based upon the same thirty-three samples used in the best conjugated diene model retained an area within that region. Wavenumbers isolated from that procedure were then able to be applied to improve the best initial model. Interferences affect model quality in two ways: those at important wavenumbers have a direct impact, while those occurring throughout the NIR spectrum can shift the outcome of the improvement process, affecting wavenumber selection and thus the quality of the optimized model. This makes presentation of representative samples to the NIR using adequate controls for environmental effects (i.e. moisture content) critical for both creation and use of models.

Models based solely on pecan oil samples were very poor while those using canola oil samples or a combined set of pecan oil samples and canola oil samples were better but not quantitative. The limited extent of oxidation in the majority of pecan oil samples was a key reason for this disparity. For all models, the bias was much smaller than the SECV, indicating that precision of the NIR results was the most significant detriment to model quality. A major contributor to the lack of precision was compositional heterogeneity within the samples. Moreover, the use of borosilicate glass lenses in the sample presentation apparatus also contributed variation at wavenumbers that may have been important for oxidation studies. The NIR instrument used transflectance in a bottom-up orientation, thus radiation had to pass through the substrate twice, once to impinge upon the sample and again to return to the detector. This not

only reduced the intensity of the beam, but because lenses varied slightly in thickness and may have seated improperly within the rotating sample holder, changes in path length may have caused spectral variations. Also, the additive used to impart resistance to thermal expansion in borosilicate glass, boron oxide, has a broad absorption centered near 1400 cm^{-1} and thus added to variation near 7000 cm^{-1} where peroxide signals have been reported.

Another issue arose from the approach of oxidizing neat oils. As peroxides are highly reactive species, this approach improved the accuracy of chemical assays by eliminating an extraction step but added the potential for conversion of analytes to other species in NIR samples upon incorporation of oxidized oils into a solid matrix. Among the best models of peroxide values, wavenumbers with the highest regression coefficients corresponded to a host of oxidized species, including aldehydes, ketones, alcohols and carboxylic acids. However, regions traditionally associated with peroxide signals were largely absent. This suggested that many of the peroxides present in oxidized oils converted to secondary oxidation products either by standard decomposition or by co-oxidation of other components of the rice flour matrix upon mixing and/or forming the NIR samples. Metal leaching from the stainless steel bowl may have catalyzed standard peroxide degradation during mixing and forming of NIR samples. Although lipid co-oxidation of proteins via radical reactions or hydroperoxide-protein amine cage reactions has been documented in model systems, it remains to be determined if these reactions occur when oil hydroperoxides are incubated within a solid food matrix.

NIR is a powerful technique that holds great promise for rapid and environmentally-sound analysis of food products. Using a simplified model solid food system and moderate amounts of sample practical for use in a research laboratory, high although not quantitative correlations for conjugated dienes and peroxide values were obtained for oils present at an amount similar to the oil content of many cereal-based and other dry food products. However, results were obtained after employing an improvement process which can be computationally intensive and is highly dependent on the quality of sample measurements. Although RPD values

for the best models herein fell at the borderline of the fair and poor classifications (see Appendix B, Table B.1), they were over twice that of the best models previously reported in similar food systems (Kaddour et al., 2006, see Table 5) based on spectral data collected with a fiber optic probe.

Improvements in sample presentation mechanisms will enhance the precision of measurements, notably the main issue degrading quality in the models developed herein. Dedicated instruments or modules can be created using optical path configurations and materials appropriate to finer analyses. Resulting improvements in spectral data will avoid hindrances to wavenumber selection techniques. Improvements in software will enable automated and systematic isolation of relevant wavenumber ranges for analytes of interest to supplant manual calculations and trial and error approaches. Together, such advances may render NIR fully effective for quantitation of lipid oxidation in solid foods in the future. Even if this goal is not ultimately achieved, these improvements will extend the range of the technology to analytes beyond the bulk properties at the core of many current applications. Prospective economic and environmental benefits of NIR indicate that further development of this technology remains a worthwhile objective.

7. FUTURE WORK

Although NIR models were unable to quantify conjugated dienes and peroxide values in oils incorporated in the model solid food systems herein, a high degree of correlation was obtained for these products and factors impacting model quality were identified. Given the potential advantages of NIR generally as well as for rapid analyses of labile products such as the indicators of lipid oxidation studied herein, further research is merited with special attention to the following areas.

The effect of sample moisture on NIR spectra should be investigated to determine whether monitoring relative humidity in the lab and/or moisture content of samples should become standard procedures when performing NIR assays of lipid oxidation. This may be necessary to distinguish contributions of water from other oxygenated products (i.e. alcohols and carboxylic acids) to certain key regions of NIR spectra. Also, one could incubate larger volumes of oil and titrate with different volumes of fresh oil to obtain samples with multiple concentrations of lipid oxidation products to assay contemporaneously. Such analysis of multiple concentrations of lipid oxidation products on the same day under identical ambient conditions would provide insight on the effect of relative humidity in lipid oxidation studies such as those herein in which samples are assayed on different days and under different ambient conditions.

The effect of sample composition should be further investigated. Larger volumes of oil can be incubated and mixed with a given sample matrix at different weight percentages (i.e. 5, 7.5, 10, 12.5 and 15% (w/w) to determine effects on spectra at each oxidation level.

The effect of co-oxidation of fats or oils of interest with constituents of the solid food matrix should be investigated. ASL studies wherein oxidized oils are mixed with an inert solid sample matrix can confirm whether peroxides in this study converted to secondary oxidation products via co-oxidation with the sample matrix. Additionally, components of the matrix used in this study (rice starch and rice protein) can be used in separate experiments to determine levels

of co-oxidation. Finally, experiments wherein oils are co-incubated with the solid matrix should be performed to determine changes in product composition and spectra.

In each case above, assays of multiple analytes should be performed contemporaneously by coordinated research teams. Otherwise, a single analyte should be assayed to ensure reference chemical and NIR assays are performed as close in time as possible and without intervening storage periods.

Finally, in conjunction with an instrument manufacturer, an improved sample presentation platform for monitoring labile analytes when using small sample volumes as well as improved software for optimizing models by selecting wavenumber ranges should be developed. In the context of lipid oxidation, sample presentation platforms should avoid the use of borosilicate glass as a sample substrate. Ideally, such platforms will protect the sample from exposure to the environment, maximize the information obtainable from a limited amount of sample, and if transflectance is employed, use a top-down orientation to avoid the need for radiation to pass through a substrate to impinge upon the sample and to return to the detector.

The above experiments and measures will be useful to address areas extrinsic to the NIR technique per se which can be improved and controlled. Success from the application of one or more of these modifications could extend the range of this technique for rapid, cost-effective and green analysis of solid product samples in the food industry as well as many other fields of endeavor.

8. REFERENCES

AOCS, AOCS Official Method Ti 1a-64: Spectrophotometric Determination of Conjugated Dienoic Acid. *INDUSTRIAL OILS AND DERIVATIVES* **2003**.

Armenta, S., Garrigues, S., & de la Guardia, M. (2007). Determination of edible oil parameters by near infrared spectrometry. *Analytica Chimica Acta, 596*, 330-337.

Babincova, M., Machova, E., & Kogan, G. (1999). Carboxymethylated glucan inhibits lipid peroxidation in liposomes. *Zeitschrift fur Naturforschung C, 54(12),* 1084-1088.

Barnes, R. J., Dhanoa, M. S., & Lister, S. J. (1989). Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Applied Spectroscopy, 43*(5), 772.

Benefits of cooking with pecan oil. (2019). Retrieved from https://pecanoil.com/pages/benefits-of-cooking-with-pecan-oil

Berzaghi, P., Dalle Zotte, A., Jansson, L. M., & Andrighetto, I. (2005). Near-infrared reflectance spectroscopy as a method to predict chemical composition of breast meat and discriminate between different n-3 feeding sources. *Poultry Science, 84*(1), 128-136.

Brereton, R. G. (2003). *Chemometrics: Data analysis for the laboratory and chemical plant*. Hoboken, NJ, USA: Wiley.

Casale, M.T., Richman, A.R., Elrod, M.J., Garland, R.M., Beaver, M.R., & Tolbert, M.A. (2007). Kinetics of acid-catalyzed aldol condensation reactions of aliphatic aldehydes. *Atmospheric Environment, 41,* 6212-6224.

Cayuela Sanchez, J. A., Moreda, W., & Garcia, J. M. (2013). Rapid determination of olive oil oxidative stability and its major quality parameters using vis/NIR transmittance spectroscopy. *Journal of Agricultural and Food Chemistry, 61*(34), 8056-8062.

Chemical compatibility database. (2019). Retrieved from https://www.coleparmer.com/chemical-resistance

Cho, S., Kim, J., & Rhee, C. (1998). Determination of rancidity of soybean oil by near infrared spectroscopy. *Journal of Near Infrared Spectroscopy, 6*(1-4), A349-A354.

Cozzolino, D., Murray, I., Chree, A., & Scaife, J. R. (2005). Multivariate determination of free fatty acids and moisture in fish oils by partial least-squares regression and near-infrared spectroscopy. *LWT--Food Science and Technology, 38*(8), 821-828.

Davies, A.M.C. (2001). Uncertainty testing in PLS regression. *Spectroscopy Europe, 13(2),* 16-19.

Dellarosa, N., Laghi, L., Martinsdóttir, E., Jónsdóttir, R., & Sveinsdóttir, K. (2015). Enrichment of convenience seafood with omega-3 and seaweed extracts: Effect on lipid oxidation. *LWT - Food Science and Technology, 62*, 746-752.

Efron, B. (1982). The jackknife estimate of variance. In Efron, B. (Ed.) *The Jackknife, the Bootstrap and Other Resampling Plans* (pp. 13-18). Philadelphia, PA: SIAM.

El-Rafey, H., Peredi, J., Kaffka, K., Naday, B., & Balogh, A. (1988). Possibilities of application of NIR technique in the analysis of oilseeds and their derivatives: Examination of used frying oils. *Fett Wissenschaft Technologie, 90*(5), 175-9.

Goddu, R. F., & Delker, D. A. (1958). Determination of terminal epoxides by near-infrared spectrophotometry. *Analytical Chemistry, 30*(12), 2013-2016.

Guo, J., Skinner, G. W., Harcum, W. W., Malone, J. P., & Weyer, L. G. (1999). Application of near-infrared spectroscopy in the pharmaceutical solid dosage form. *Drug Development and Industrial Pharmacy, 25*(12), 1267-1270.

Hanst, P. L., Early, V. H., & Klemperer, W. (1965). Infrared spectrum and molecular structure of B2O3. *The Journal of Chemical Physics, 42*(3), 1097-1104.

Holman, R. T., & Edmondson, P. R. (1956). Near-infrared spectra of fatty acids and some related substances. *Analytical Chemistry, 28*(10), 1533-1538.

Holman, R., Nickell, C., Privett, O., & Edmondson, P. (1958). Detection and measurement of hydroperoxides by near infrared spectrophotometry. *Journal of the American Oil Chemists' Society (JAOCS), 35*(8), 422.

Hong, J.H., Koseki, S.Y., & Yasumoto, K. (1994). Analysis of peroxide values in edible oils by near-infrared spectroscopy. *Nippon Shokuhin Kogyo Gakkaishi, 41,* 277-280.

Jensen, P. N., Christensen, J., & Engelsen, S. B. (2004). Oxidative changes in pork scratchings, peanuts, oatmeal and muesli viewed by fluorescence, near-infrared and infrared spectroscopy. *European Food Research and Technology, 219*(3), 294-304.

Jensen, P. N., Danielsen, B., Bertelsen, G., Skibsted, L. H., & and Andersen, M. L. (2005). Storage stabilities of pork scratchings, peanuts, oatmeal and muesli: Comparison of ESR spectroscopy, headspace-GC and sensory evaluation for detection of oxidation in dry foods. *Food Chemistry, 91*, 25-38.

Jensen, P. N., Sorensen, G., Engelsen, S. B., & Bertelsen, G. (2001). Evaluation of quality changes in walnut kernels (*Juglans regia* L.) by vis/NIR spectroscopy. *Journal of Agricultural and Food Chemistry, 49*(12), 5790-5796.

Kaddour, A. A., Grand, E., Barouh, N., Barea, B., Villeneuve, P., & Cuq, B. (2006). Near-infrared spectroscopy for the determination of lipid oxidation in cereal food products. *European Journal of Lipid Science and Technology, 108*(12), 1037-1046.

Karlsdottir, M. G., Arason, S., Kristinsson, H. G., & Sveinsdottir, K. (2014). The application of near infrared spectroscopy to study lipid characteristics and deterioration of frozen lean fish muscles. *Food Chemistry, 159*, 420-427.

Klein, R. A. (1970). The detection of oxidation in liposome preparations. *Biochim. Biophys. Acta, 210,* 486-489.

Kostik, V., Memeti, S., & Bauer, B. (2013). Fatty acid composition of edible oils and fats. *Journal of Hygienic Engineering and Design, 4*, 112-116.

Mandal, T. K., Ghose, S., Sur, P., & Chatterjee, S. N. (1978). Effect of ultra-violet radiation on the liposomal membrane. *International Journal of Radiation Biology and Related Studies in Physics, Chemistry and Medicine, 33*(1), 75-79.

Manley, M., & Eberle, K. (2006). Comparison of Fourier transform near infrared spectroscopy partial least square regression models for South African extra virgin olive oil using spectra collected on two spectrophotometers at different resolutions and path lengths. *Journal of Near Infrared Spectroscopy, 14*(2), 111-126.

Martens, H., & Martens, M. (2000). Modified jack-knife estimation of parameter uncertainty in bilinear modelling by partial least squares regression (PLSR). *Food Quality and Preference, 11(1),* 5-16.

Naes, T., Isaksson, T., Fearn, T., & Davies, T. (2002). *A user-friendly guide to multivariate calibration and classification.* Chichester, UK: NIR Publications.

NIRCal 5.4 software manual. New Castle, Delaware, USA: Buchi Corporation.

Osborne, B. G., & Fearn T. (1986). *Near infrared spectroscopy in food analysis*. New York, NY, USA: Longman Scientific & Technical / John Wiley & Sons, Inc.

Park, T. S., Bae, Y. M., Seo, H. S., Park, T. J., Seol, K. H., Lim, D. K., et al. (2008). Evaluation of pork loin freshness using absorbance characteristic of near-infrared. *Biological Engineering, 1*(2), 173-180.

Peck, M. C. P., Carter, R. O., & Qaderi, S. B. A. (1987). Near infrared measurements of terminal epoxides in polymer resin systems. I. Analytical considerations. *Journal of Applied Polymer Science, 33*(1), 77-86.

Pellegrini, N., Serafini, M., Colombi, B., Del Rio, D., Salvatore, S., Bianchi, M., & Brighenti, F. (2003). Total antioxidant capacity of plant foods, beverages and oils consumed in Italy assessed by three different in vitro assays. *The Journal of Nutrition, 133*(9), 2812-2819.

Realini, C. E., Duckett, S. K., & Windham, W. R. (2004). Effect of vitamin C addition to ground beef from grass-fed or grain-fed sources on color and lipid stability, and prediction of fatty acid composition by near-infrared reflectance analysis. *Meat Science, 68*(1), 35-43.

Roggo, Y., Jent, N., Edmond, A., Chalus, P., & Ulmschneider, M. (2005). Characterizing process effects on pharmaceutical solid forms using near-infrared spectroscopy and infrared imaging. *European Journal of Pharmaceutics and Biopharmaceutics, 61,* 100-110.

Savitzky, A., & Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry, 36*(8), 1627-1639.

Schaich, K.M. & Karel, M. (1976). Free radical reactions of peroxidizing lipids with amino acids and proteins: An ESR study. *Lipids, 11,* 392-400.

Schaich, K.M. (2005). Lipid Oxidation: Theoretical Aspects. In Fereidoon Shahidi (Ed.) *Bailey's Industrial Oil and Fat Products* (Sixth ed., pp. 269-355). Hoboken, NJ: John Wiley & Sons, Inc.

Schaich, K.M. (2008). Co-oxidations of oxidizing lipids: Reactions with proteins. In Kamal-Eldin, A., Min, D.B. (Eds.) *Lipid Oxidation Pathways, Vol.2* (pp. 183-274). Boca Raton, FL: CRC Press.

Siesler, H.W., Ozaki, Y., Kawata, S., & Heise, H.M. (2002). Appendix: Group frequencies in the Near-Infrared region. In Siesler, H.W., Ozaki, Y., Kawata, S., & Heise, H.M. (Eds.) *Near-Infrared Spectroscopy: Principles, Instruments, Applications* (pp. 335-339). Weinheim, Germany: Wiley-VCH.

Stainless steel 304 - alloy composition. (2019). Retrieved from http://www.espimetals.com/index.php/190-technical-data/stainless-steel-304-alloy-composition/200-stainless-steel-304-alloy-composition

Szabo, A., Bazar, G., Andrassy-Baka, G., Locsmandi, L., & Romvari, R. (2009). A near infrared spectroscopic (NIR) approach to estimate quality alterations during prolonged heating of lard. *Acta Alimentaria, 38*(1), 97-106.

Takamura, H., Hyakumoto, N., Endo, N., Matoba, T., & Nishiike, T. (1995). Determination of lipid oxidation in edible oils by near infrared spectroscopy. *Journal of Near Infrared Spectroscopy, 3*(4), 219-225.

Toro-Vazquez, J. F., Charo-Alonso, M. A., & Perez-Briceno, F. (1999). Fatty acid composition and its relationship with physicochemical properties of pecan (*Carya illinoensis*) oil. *Journal of the American Oil Chemists' Society, 76*(8), 957-965.

Tukey, J.W. (1958). Bias and confidence in not-quite large samples (Abstract). *The Annals of Mathematical Statistics, 29,* 614.

Vieira, F. S., & Pasquini, C. (2013). Near infrared emission photometer for measuring the oxidative stability of edible oils. *Analytica Chimica Acta, 796*, 101-107.

Westad, F., & Martens, H. (2000). Variable selection in near infrared spectroscopy based on significance testing in partial least squares regression. *Journal of Near Infrared Spectroscopy, 8*(2), 117-124.

Williams, P. C. (2001). Implementation of near-infrared technology. In P. Williams, & K. Norris (Eds.), *Near-infrared technology in the agricultural and food industries* (Second ed., pp. 145-169). St. Paul, Minnesota, USA: American Association of Cereal Chemists, Inc.

Wójcicki, K., Khmelinskii, I., Sikorski, M., & Sikorska, E. (2015). Near and mid infrared spectroscopy and multivariate data analysis in studies of oxidation of edible oils. *Food Chemistry, 187*, 416-423.

Workman, J., & Weyer L. (2008). *Practical guide to interpretive near-infrared spectroscopy*. New York, NY, USA: CRC Press.

Wu, W., Walczak, B., Massart, D. L., Prebble, K. A., & Last, I. R. (1995). Spectral transformation and wavelength selection in near-infrared spectra classification. *Analytica Chimica Acta, 315*(3), 243-255.

Wu, X., Beecher, G. R., Holden, J. M., Haytowitz, D. B., Gebhardt, S. E., & Prior, R. L. (2004). Lipophilic and hydrophilic antioxidant capacities of common foods in the United States. *Journal of Agricultural and Food Chemistry, 52*(12), 4026-4037.

Yao, L. (2015). Development of new methods for analyzing lipid oxidation. (PhD, Rutgers, The State University of New Jersey).

Yildiz, G., Wehling, R. L., & Cuppett, S. L. (2001). Method for determining oxidation of vegetable oils by near-infrared spectroscopy. *Journal of the American Oil Chemists' Society, 78*(5), 495-502.

Yildiz, G., Wehling, R. L., & Cuppett, S. L. (2003). Comparison of four analytical methods for the determination of peroxide value in oxidized soybean oils. *Journal of the American Oil Chemists' Society, 80*(2), 103-107.

Yoon, J., Kim, J., Duy, P. K., Kim, M., Kang, S., & Chung, H. (2013). Diffuser-incorporated transmission NIR measurement for reliable analysis of packed granular samples. *Analyst, 138,* 4922-4932.

APPENDIX A

Colthup Chart from NIR Studies of Lipids and Lipid Oxidation

[Adapted from NIRCal 5.4 Software Module, Appendix 4a of Workman & Weyer (2008) and Lipid Oxidation Literature]

Wavenumber (cm^{-1})	Feature	Compound or Lipid System	Reference
1744	Carbonyl (tentative; slight change)	Olive; Sunflower; Rapeseed Oil	Wojcicki et al., 2015
3007	Cis Unsaturation CH (small)	Pork Scratchings; Peanuts; Oatmeal; Muesli	Jensen et al., 2004
3448	Hydroperoxides	Rapeseed; Corn; Soybean; Sunflower; Peanut; Extra Virgin Olive Oil	Vieira & Pasquini, 2013
3500	Peroxides	Olive; Sunflower; Rapeseed	Wojcicki et al., 2015
3534	Carboxylic Acid (intense)	Standards	Holman & Edmondson, 1956
3546	Hydroperoxide (tentative)	Fatty Acid Standards	Holman et al., 1958
3610	Hydroxyl	Fatty Acid Standards	Holman et al., 1958
3636	Alcohol Hydroxyl (intense)	Standards	Holman & Edmondson, 1956
3676	Carboxylic Acid (intense)	Standards	Holman & Edmondson, 1956
4000	CH	Standards	Holman & Edmondson, 1956
4000	CH Stretch, CC and COC Stretch Combination	Polysaccharides	Workman & Weyer, 2008
4019	CH Stretch and CC Stretch Combination	Cellulose	Workman & Weyer, 2008
4049	Methylene CH Combination	Lipids and Aliphatic Compounds	Workman & Weyer, 2008
4049	Amide	Proteins	Workman & Weyer, 2008
4052	Protein	Marine & Linseed Oil Dietary Supplements	Berzaghi et al., 2005
4060	Amide	Proteins	Workman & Weyer, 2008
4063	C-H Str.+ C-C Str.	Starch	NIRCal 5.4
4068	Methyl CH	Aliphatic Compounds	Workman & Weyer, 2008

Wavenumber (cm^{-1})	Feature	Compound or Lipid System	Reference
4090	Amide	Proteins	Workman & Weyer, 2008
4202	O-H def. Second overtone	ROH	NIRCal 5.4
4202	CH Stretch and CC Stretch Combination	Lipids	Workman & Weyer, 2008
4232	Methylene CH Combination	Aliphatic Compounds (Branched)	Workman & Weyer, 2008
4252	C-H def. Second overtone	Cellulose	NIRCal 5.4
4252	CH Bending	Polysaccharides	Workman & Weyer, 2008
4255	Oil	Ground Beef from Grass or Grain Fed Cattle	Realini et al., 2004
4259	CH	Soybean Oil	Cho et al., 1998
4259	CH	Fish Oil	Cozzolino et al., 2005
4261	CH2 sym. Str.+ =CH2 def.	HC=CHCH2	NIRCal 5.4
4261	Methylene CH Combination	Aliphatic Compounds (Linear)	Workman & Weyer, 2008
4264	CH2 Stretch and Bend Combination	Extra Virgin Olive Oil	Manley & Eberle, 2006
4274	CH2	Standards	Holman & Edmondson, 1956
4274	Part of Conjugated Triene Triplet (also slightly evident in t,t-Conjugated Dienes)	Standards	Holman & Edmondson, 1956
4274	CH Combination	Lean Fish	Karlsdottir et al., 2014
4281	C-H Str.+ C-H def.	Cellulose	NIRCal 5.4
4281		Cellulose	Workman & Weyer, 2008
4283	CH Stretch and CH2 Deformation Combination	Polysaccharides	Workman & Weyer, 2008
4292	CH Stretch and CH2 Deformation Combination	Polysaccharides	Workman & Weyer, 2008
4305	C-H Str.+ C-H def.	CH2	NIRCal 5.4

Wavenumber (cm^{-1})	Feature	Compound or Lipid System	Reference
4307	CH Stretch and CH2 Deformation Combination	Polysaccharides	Workman & Weyer, 2008
4310	Part of Conjugated Triene Triplet (also slightly evident in t,t-Conjugated Dienes)	Standards	Holman & Edmondson, 1956
4314	Methylene	Aliphatic Compounds (Branched)	Workman & Weyer, 2008
4329	CH2 Stretch and Bend Combination	Extra Virgin Olive Oil	Manley & Eberle, 2006
4329	N-H Str.+ C-H def.	CH2	NIRCal 5.4
4329	CHO	Lipids	Workman & Weyer, 2008
4329	CH Bending	Lipids	Workman & Weyer, 2008
4333	CH	Fish Oil	Cozzolino et al., 2005
4333	Methylene	Aliphatic Compounds (Linear)	Workman & Weyer, 2008
4337	CH	Soybean Oil	Cho et al., 1998
4337	Oil	Ground Beef from Grass or Grain Fed Cattle	Realini et al., 2004
4348	CH2	Standards	Holman & Edmondson, 1956
4348	Part of Conjugated Triene Triplet (also slightly evident in t,t-Conjugated Dienes)	Standards	Holman & Edmondson, 1956
4348	CH2 (more intense for fully saturated fatty acids)	Palm; Rapeseed	Kaddour et al., 2006
4348	CH Combination	Lean Fish	Karlsdottir et al., 2014
4348	CH Bending	Amides	Workman & Weyer, 2008
4359	N-H Str.+ C=O Str.	Amino Acid	NIRCal 5.4
4365-4370	Amide	Proteins	Workman & Weyer, 2008
4383	C-H Str.+ C-H def.	CH3	NIRCal 5.4

Wavenumber (cm^{-1})	Feature	Compound or Lipid System	Reference
4386	CH3	Standards	Holman & Edmondson, 1956
4386	CH Stretch and Methylene Deformation	Starch	Workman & Weyer, 2008
4394	O-H Str.+ C-C Str.	Starch	NIRCal 5.4
4400	OH Stretch and CO Stretch Combination	Glucose	Workman & Weyer, 2008
4405	Amide	Proteins	Workman & Weyer, 2008
4405	OH Stretch and CO Stretch Combination	Cellulose	Workman & Weyer, 2008
4425	CH3	Standards	Holman & Edmondson, 1956
4425	Methyl Ester (sharp)	Standards	Holman & Edmondson, 1956
4440	O-H Str.+ O-H def.	Starch	NIRCal 5.4
4444	Protein	Lean Fish	Karlsdottir et al., 2014
4456	Canola, Olive and Safflower Oils Characteristic Absorbance (possible secondary oxidation product)	Canola; Olive; Safflower; Soybean; Cottonseed; Methyl Oleate; Methyl Linoleate	Takamura et al., 1995
4460	N-H Str.+ NH3+ def.	Amino Acid	NIRCal 5.4
4520-4570	CH Stretch and C=O Stretch Combination	Aldehydes	Siesler, et al., 2002
4525	Aldehyde Carbonyl (weak)	Standards	Holman & Edmondson, 1956
4525-4540	Amide	Proteins	Workman & Weyer, 2008
4532	Terminal Epoxides	Standards	Peck et al., 1987
4545	C-H Str.+ C=O Str.	-CHO	NIRCal 5.4
4545	CH Stretch and C=O Combination	Carbohydrates	Workman & Weyer, 2008
4545	CH	Standards	Holman & Edmondson, 1956

Wavenumber (cm^{-1})	Feature	Compound or Lipid System	Reference
4545	Deviation in Hoki Spectra related to Oxygenated Compounds	Lean Fish	Karlsdottir et al., 2014
4545	CHO	Aldehyde	Osborne & Fearn, 1986
4545	Terminal Epoxides (intense)	Standards	Goddu & Delker, 1958
4550	Linoleic Acid Content Determination of oil Heated at 200-210°C	Sunflower Oil (Fried)	El-Rafey et al., 1988
4550-5550	Broad OH Band	Water, Alcohols or Polyols	Workman & Weyer, 2008
4566	CH2 asym. Str.+ C= Str.	HC=CH	NIRCal 5.4
4566	Cis Unsaturation	Standards	Holman & Edmondson, 1956
4566	Cis Unsaturation	Fatty Acid Standards	Holman et al., 1958
4566	Fatty Acids (Cis Unsaturation)	Marine & Linseed Oil Dietary Supplements	Berzaghi et al., 2005
4587	2x amide I+ amide III	Protein	NIRCal 5.4
4587	N-H	Proteins	Workman & Weyer, 2008
4587	Polar Content Determination of oil Heated at 200-210°C	Sunflower Oil (Fried)	El-Rafey et al., 1988
4590	-C=C- Combination (decrease)	Olive; Sunflower; Rapeseed	Wojcicki et al., 2015
4596	Polar Content Determination of oil Heated at 170-180°C	Sunflower Oil (Fried)	El-Rafey et al., 1988
4596	Polymer Content Determination of oil Heated at 170-180°C	Sunflower Oil (Fried)	El-Rafey et al., 1988
4600	Protein	Marine & Linseed Oil Dietary Supplements	Berzaghi et al., 2005
4604	Linoleic Acid Content Determination of oil Heated at 200-210°C	Sunflower Oil (Fried)	El-Rafey et al., 1988

Wavenumber (cm^{-1})	Feature	Compound or Lipid System	Reference
4608	CH of HC=CH	Alkenes	Workman & Weyer, 2008
4615	Amide	Proteins	Workman & Weyer, 2008
4621	Linoleic Acid Content Determination of oil Heated at 200-210°C	Sunflower Oil (Fried)	El-Rafey et al., 1988
4630	2x amide I+ amide III	CONHR	NIRCal 5.4
4638	Polymer Content Determination of oil Heated at 170-180°C	Sunflower Oil (Fried)	El-Rafey et al., 1988
4651	2x amide I+ amide III	CONH2	NIRCal 5.4
4651	Cis Unsaturation	Standards	Holman & Edmondson, 1956
4651	Cis Unsaturation	Fatty Acid Standards	Holman et al., 1958
4651	Fatty Acids	Marine & Linseed Oil Dietary Supplements	Berzaghi et al., 2005
4664	Cis Unsaturation and CH Combination	Fish Oil	Cozzolino et al., 2005
4670	-C=C- Combination (decrease)	Olive; Sunflower; Rapeseed	Wojcicki et al., 2015
4673	Methyl Ester (sharp)	Standards	Holman & Edmondson, 1956
4673	=C-H Str.+ C=C Str.	HC=CH	NIRCal 5.4
4673	CH Stretch and C=O Stretch Combination	Lipids	Workman & Weyer, 2008
4690	N-H Str.+ C=O Str.	Amino Acid	NIRCal 5.4
4739	N-H sym. Str.+ amide III	CONH2,CONHR	NIRCal 5.4
4762	Cis Unsaturation and CH Combination for Terminal Double Bonds (Characteristic of Fish Oil)	Fish Oil	Cozzolino et al., 2005
4762	Polymer Content Determination of oil Heated at 200-210°C	Sunflower Oil (Fried)	El-Rafey et al., 1988

Wavenumber (cm^{-1})	Feature	Compound or Lipid System	Reference
4762	2x O-H def.+ 2x C-O Str.	Starch	NIRCal 5.4
4762	Carbohydrates	Carbohydrates	Workman & Weyer, 2008
4770	OH Deformation	Water or Alcohols	Workman & Weyer, 2008
4798	PV (high correlation)	Canola; Olive; Safflower; Soybean; Cottonseed; Methyl Oleate; Methyl Linoleate	Takamura et al., 1995
4798	Canola, Olive and Safflower Oils Characteristic Absorbance (possible secondary oxidation product); PV in Soya bean and Cottonseed (low PV range oils)	Canola; Olive; Safflower; Soybean; Cottonseed; Methyl Oleate; Methyl Linoleate	Takamura et al., 1995
4804	Peroxides (increase)	Olive; Sunflower; Rapeseed	Wojcicki et al., 2015
4808	PV Calibration Wavelength	Soybean Oil	Cho et al., 1998
4808	Hydroperoxide	Standards	Holman & Edmondson, 1956
4808	Linoleic Acid Content Determination of oil Heated at 170-180°C	Sunflower Oil (Fried)	El-Rafey et al., 1988
4808	O-H Str.+ O-H def.	ROH, Sucrose, Starch	NIRCal 5.4
4810	Hydroperoxides	Olive; Sunflower; Rapeseed	Wojcicki et al., 2015
4831	Alcohol Hydroxyl (weak)	Standards	Holman & Edmondson, 1956
4831	Hydroperoxide	Fatty Acid Standards	Holman et al., 1958
4831	PV (significant correlation)	Soybean Oil	Yildiz et al., 2001
4850	NH Combination (Secondary Amides)	Proteins	Workman & Weyer, 2008
4854	Amide	Proteins	Workman & Weyer, 2008
4864	Protein	Marine & Linseed Oil Dietary Supplements	Berzaghi et al., 2005

Wavenumber (cm^{-1})	Feature	Compound or Lipid System	Reference
4865	Amide	Proteins	Workman & Weyer, 2008
4866	Amide NH Stretch and C=O Stretch Combination	Proteins	Workman & Weyer, 2008
4873	Linoleic Acid Content Determination of oil Heated at 170-180°C	Sunflower Oil (Fried)	El-Rafey et al., 1988
4878	N-H Str.+ amide II	CONH2, Protein	NIRCal 5.4
4925	NH and CN Combination	Primary Amides	Workman & Weyer, 2008
4926	C=O Str. Second overtone	CONH2	NIRCal 5.4
4926	C=O; Acid Value Estimation	Lard	Szabo et al., 2009
4926	Amine	Standards	Holman & Edmondson, 1956
4950	PV Calibration Wavelength	Soybean Oil	Cho et al., 1998
4975	NH and CN Combination	Primary Amides	Workman & Weyer, 2008
4975	Amine	Standards	Holman & Edmondson, 1956
4975	Hydroxyl	Fatty Acid Standards	Holman et al., 1958
4980	C=O from liberated fatty acids	Soybean Oil	Cho et al., 1998
5000	2x O-H def.+ C-O def.	Starch	NIRCal 5.4
5000	N-H sym. Str.+ amide II	CONH2,CONHR	NIRCal 5.4
5051	N-H asym. Str.+ amide II	Protein	NIRCal 5.4
5051	Amide	Proteins	Workman & Weyer, 2008
5076	Linoleic Acid Content Determination of oil Heated at 170-180°C	Sunflower Oil (Fried)	El-Rafey et al., 1988
5100	NH Combination	Primary Amides	Workman & Weyer, 2008
5102	N-H asym.Str.+ amide II	CONH2	NIRCal 5.4
5102	Polymeric OH	Polysaccharides	Workman & Weyer, 2008

Wavenumber (cm^{-1})	Feature	Compound or Lipid System	Reference
5112	Linoleic Acid Content Determination of oil Heated at 170-180°C	Sunflower Oil (Fried)	El-Rafey et al., 1988
5128	C=O Str. Second overtone	-CO2R	NIRCal 5.4
5128	C=O / C=OOR	Acids and Esters	Workman & Weyer, 2008
5128	Methyl Ester (weak)	Standards	Holman & Edmondson, 1956
5155	O-H Str.+ O-H def.	H2O	NIRCal 5.4
5155	OH Stretch and HOH Bending Combination	Water	Workman & Weyer, 2008
5165	Water	Ground Beef from Grass or Grain Fed Cattle	Realini et al., 2004
5181	OH Stretch and HOH Bending Combination	Polysaccharides	Workman & Weyer, 2008
5181	Moisture	Marine & Linseed Oil Dietary Supplements	Berzaghi et al., 2005
5187	OH (Hydroxyl)	Fish Oil	Cozzolino et al., 2005
5195	OH Water Combination	Palm; Rapeseed	Kaddour et al., 2006
5200	OH Stretch and HOH Deformation Combination	Water	Workman & Weyer, 2008
5208	C=O Str. Second overtone	CONH	NIRCal 5.4
5208	Amide	Amides	Workman & Weyer, 2008
5208	Water	Lean Fish	Karlsdottir et al., 2014
5210-5290	2x C=O Stretch	Carboxylic Acids	Siesler, et al., 2002
5219	Canola, Olive and Safflower Oils Characteristic Absorbance (possible secondary oxidation product)	Canola; Olive; Safflower; Soybean; Cottonseed; Methyl Oleate; Methyl Linoleate	Takamura et al., 1995

Wavenumber (cm^{-1})	Feature	Compound or Lipid System	Reference
5241	O-H Str. first overtone	ROH	NIRCal 5.4
5263	C=O Str. Second overtone	-CO2H	NIRCal 5.4
5263	O-H Str.+2x C-O Str.	Starch	NIRCal 5.4
5263	Carboxylic Acid (intensity drops with increasing chain length)	Standards	Holman & Edmondson, 1956
5263	Methyl Ester (weak)	Standards	Holman & Edmondson, 1956
5400	Polymer Content Determination of oil Heated at 200-210°C	Sunflower Oil (Fried)	El-Rafey et al., 1988
5405	Polymer Content Determination of oil Heated at 200-210°C	Sunflower Oil (Fried)	El-Rafey et al., 1988
5459	CH (Carbonyls)	Fish Oil	Cozzolino et al., 2005
5495	O-H Str.+2x C-O Str.	Cellulose	NIRCal 5.4
5495	OH Stretch and CO Stretch Combination	Cellulose	Workman & Weyer, 2008
5543	Polymer Content Determination of oil Heated at 200-210°C	Sunflower Oil (Fried)	El-Rafey et al., 1988
5587	OH Combination	Water	Workman & Weyer, 2008
5618	C-H Str. first overtone	Cellulose	NIRCal 5.4
5618	Methylene CH	Cellulose	Workman & Weyer, 2008
5631	Linoleic Acid Content Determination of oil Heated at 170-180°C	Sunflower Oil (Fried)	El-Rafey et al., 1988
5650	CH2 in Acids or Alcohols	Standards	Holman & Edmondson, 1956
5666	CH2	Extra Virgin Olive Oil	Manley & Eberle, 2006
5666	C-H Str. first overtone	CH2	NIRCal 5.4
5666	Methylene CH (Asymmetric)	Aliphatic Compounds	Workman & Weyer, 2008
5675	Methylene CH (Symmetric)	Aliphatic Compounds	Workman & Weyer, 2008

Wavenumber (cm^{-1})	Feature	Compound or Lipid System	Reference
5675	Oil	Ground Beef from Grass or Grain Fed Cattle	Realini et al., 2004
5682	CH	Soybean Oil	Cho et al., 1998
5682	CH2 (more intense for fully saturated fatty acids)	Palm; Rapeseed	Kaddour et al., 2006
5682	CH2	Marine & Linseed Oil Dietary Supplements	Berzaghi et al., 2005
5682	CH	Fish Oil	Cozzolino et al., 2005
5682	CH	Lean Fish	Karlsdottir et al., 2014
5708	Acid Value Calibration Wavelength	Soybean Oil	Cho et al., 1998
5734	Canola, Olive and Safflower Oils Characteristic Absorbance; PV (correlation)	Canola; Olive; Safflower; Soybean; Cottonseed; Methyl Oleate; Methyl Linoleate	Takamura et al., 1995
5747	S-H Str. first overtone	-SH	NIRCal 5.4
5747	SH	Thiols	Workman & Weyer, 2008
5747	CH2 in Acids or Alcohols	Standards	Holman & Edmondson, 1956
5755	Amide	Proteins	Workman & Weyer, 2008
5765	Methyl CH (NH2CH3)	Amines	Workman & Weyer, 2008
5767	CH	Olive; Maize; Seed; Sunflower	Armenta et al., 2007
5767	CH2	Extra Virgin Olive Oil	Manley & Eberle, 2006
5770	Methyl CH (ROCH3)	Ethers	Workman & Weyer, 2008
5773	Methyl CH (ROHCH3)	Alcohols	Workman & Weyer, 2008
5780	Oil	Ground Beef from Grass or Grain Fed Cattle	Realini et al., 2004
5787	Methylene CH (Asymmetric)	Aliphatic Compounds	Workman & Weyer, 2008
5797	C-H Str. first overtone	CH2	NIRCal 5.4
5797	Methylene CH	Aliphatic Compounds	Workman & Weyer, 2008

Wavenumber (cm^{-1})	Feature	Compound or Lipid System	Reference
5797	CH2 (more intense for fully saturated fatty acids)	Palm; Rapeseed	Kaddour et al., 2006
5807	CH	Soybean Oil	Cho et al., 1998
5807	CH2	Marine & Linseed Oil Dietary Supplements	Berzaghi et al., 2005
5814	CH	Fish Oil	Cozzolino et al., 2005
5829	CH	Olive; Maize; Seed; Sunflower	Armenta et al., 2007
5853	Terminal Methyl CH (RCH3)	Aliphatic Compounds	Workman & Weyer, 2008
5865	C-H Str. first overtone	CH3	NIRCal 5.4
5865	Methyl CH	Aliphatic Compounds	Workman & Weyer, 2008
5869	Linoleic Acid Content Determination of oil Heated at 170-180°C	Sunflower Oil (Fried)	El-Rafey et al., 1988
5869	Methyl CH	Aliphatic Compounds	Workman & Weyer, 2008
5870	CH2 (decrease)	Olive; Sunflower; Rapeseed	Wojcicki et al., 2015
5872	Methyl CH	Aliphatic Compounds (Branched)	Workman & Weyer, 2008
5880	Methyl CH (ROCH3)	Ethers	Workman & Weyer, 2008
5880	Methyl CH (ROHCH3)	Alcohols	Workman & Weyer, 2008
5882	CH Stretch	Soybean Oil	Yildiz et al., 2001
5882	- CH3	Marine & Linseed Oil Dietary Supplements	Berzaghi et al., 2005
5882	Characteristic of Oxygenated Products of Lipid Oxidation in Fish Oils	Fish Oil	Cozzolino et al., 2005
5882	CH	Standards	Holman & Edmondson, 1956
5882	CH3 in Acids or Alcohols	Standards	Holman & Edmondson, 1956
5898	Methyl CH Adjacent to Carbonyl	Methyl Ketones	Workman & Weyer, 2008

Wavenumber (cm^{-1})	Feature	Compound or Lipid System	Reference
5900	C-H Str. first overtone	CH3	NIRCal 5.4
5900	Methyl CH	Aliphatic Compounds	Workman & Weyer, 2008
5903	Methyl CH (Asymmetric)	Aliphatic Compounds	Workman & Weyer, 2008
5905	Methyl CH	Aliphatic Compounds	Workman & Weyer, 2008
5908	Methyl CH Beta to Carbonyl	Ethyl Ketones	Workman & Weyer, 2008
5915-5925	Amide	Proteins	Workman & Weyer, 2008
5925	Methyl CH (ROHCH3)	Alcohols	Workman & Weyer, 2008
5935	C-H Str. first overtone	Aromatic	NIRCal 5.4
5938	Canola, Olive and Safflower Oils Characteristic Absorbance (possible secondary oxidation product)	Canola; Olive; Safflower; Soybean; Cottonseed; Methyl Oleate; Methyl Linoleate	Takamura et al., 1995
5946	Methyl CH Beta to Carbonyl	Ethyl Ketones	Workman & Weyer, 2008
5952	Fatty Acids	Marine & Linseed Oil Dietary Supplements	Berzaghi et al., 2005
5952	CH	Lean Fish	Karlsdottir et al., 2014
5960	Methyl CH Adjacent to Carbonyl	Methyl Ketones	Workman & Weyer, 2008
6010	Polar Content Determination of oil Heated at 200-210°C	Sunflower Oil (Fried)	El-Rafey et al., 1988
6011	Methyl CH (ROHCH3)	Alcohols	Workman & Weyer, 2008
6024	C-H Str. first overtone	cis-RCH=CHR`	NIRCal 5.4
6039	Linoleic Acid Content Determination of oil Heated at 200-210°C	Sunflower Oil (Fried)	El-Rafey et al., 1988
6053	Polar Content Determination of oil Heated at 170-180°C	Sunflower Oil (Fried)	El-Rafey et al., 1988
6060-6100	CH Stretch 1st Overtone	Epoxides	Siesler, et al., 2002

Wavenumber (cm^{-1})	Feature	Compound or Lipid System	Reference
6061	Terminal Epoxides	Standards	Goddu & Delker, 1958
6079	C-H Str. first overtone	R-CH-CH	NIRCal 5.4
6120	Vinyl CH (CH2=CH-)	Aliphatic Compounds	Workman & Weyer, 2008
6143	Polymer Content Determination of oil Heated at 170-180°C	Sunflower Oil (Fried)	El-Rafey et al., 1988
6173	C-H Str. first overtone	=CH2	NIRCal 5.4
6173	Alkene CH	Alkenes	Workman & Weyer, 2008
6200	Vinyl CH Conjugated to Carbonyl	Alkenes	Workman & Weyer, 2008
6240-6850	Hydrogen Bonded Alcohol OH	Alcohols	Workman & Weyer, 2008
6240-7100	OH Stretch	Water or Alcohols	Workman & Weyer, 2008
6250-6540	NH (Secondary Amide)	Proteins	Workman & Weyer, 2008
6319	OH Stretch	Water or Alcohols	Workman & Weyer, 2008
6329	O-H Str. first overtone (intermol.H-bond)	Glucose	NIRCal 5.4
6330	OH Combination (Broad)	Water or Alcohols	Workman & Weyer, 2008
6369	N-H Str. first overtone	-CONH-	NIRCal 5.4
6369	Amide	Proteins	Workman & Weyer, 2008
6369	NH	Palm; Rapeseed	Kaddour et al., 2006
6471	NH (Secondary Amine)	Secondary Amines	Workman & Weyer, 2008
6494	O-H Str. first overtone (intramol.H-bond)	Starch	NIRCal 5.4
6494	Polymeric OH	Starch / Polymeric Alcohols	Workman & Weyer, 2008
6523	C-H Str. first overtone	=CH	NIRCal 5.4
6536	Amine	Standards	Holman & Edmondson, 1956
6536	N-H Str. first overtone	RNH2	NIRCal 5.4
6536	Methyne CH		Workman & Weyer, 2008
6536	Amide	Proteins	Workman & Weyer, 2008

Wavenumber (cm^{-1})	Feature	Compound or Lipid System	Reference
6536	NH (Secondary Amine)	Secondary Amines	Workman & Weyer, 2008
6545	O-H Str. first overtone (intramol.H-bond)	Starch	NIRCal 5.4
6579	N-H Str. first overtone (intramol.H-bond)	CONH2	NIRCal 5.4
6579	O-H Str. first overtone	ROH	NIRCal 5.4
6579	Amide	Proteins	Workman & Weyer, 2008
6580	NH (Secondary Amine)	Secondary Amines	Workman & Weyer, 2008
6596	Protein	Marine & Linseed Oil Dietary Supplements	Berzaghi et al., 2005
6623	N-H Str. first overtone	Proteins	NIRCal 5.4
6623	Amide	Proteins	Workman & Weyer, 2008
6667	N-H Str. first overtone	NH	NIRCal 5.4
6667	Amide	Proteins	Workman & Weyer, 2008
6702	N-H Str. first overtone	ArNH2	NIRCal 5.4
6702	Amide	Proteins	Workman & Weyer, 2008
6710	NH (Primary Amides)	Primary Amides	Workman & Weyer, 2008
6711	N-H Str. first overtone	CONHR	NIRCal 5.4
6711	O-H Str. first overtone (intramol.H-bond)	Cellulose	NIRCal 5.4
6711	NH (Secondary Amide)	Proteins	Workman & Weyer, 2008
6711	Polymeric OH	Starch / Polymeric Alcohols	Workman & Weyer, 2008
6729	NH (Secondary Amine)	Secondary Amines	Workman & Weyer, 2008
6743	N-H Str. first overtone	CONH2	NIRCal 5.4
6743	Amide	Proteins	Workman & Weyer, 2008
6751	NH (Secondary Amine)	Secondary Amines	Workman & Weyer, 2008

Wavenumber (cm^{-1})	Feature	Compound or Lipid System	Reference
6757	O-H Str. first overtone (intramol.H-bond)		NIRCal 5.4
6798	N-H Str. first overtone	CONHR	NIRCal 5.4
6798	NH (Secondary Amide)	Proteins	Workman & Weyer, 2008
6805	NH (Primary Amides)	Primary Amides	Workman & Weyer, 2008
6812	Canola, Olive and Safflower Oils Characteristic Absorbance (possible secondary oxidation product)	Canola; Olive; Safflower; Soybean; Cottonseed; Methyl Oleate; Methyl Linoleate	Takamura et al., 1995
6812	Tertiary Alcohol OH		Workman & Weyer, 2008
6826	NH (Secondary Amine)	Secondary Amines	Workman & Weyer, 2008
6831	Secondary Alcohol OH		Workman & Weyer, 2008
6835	Amide	Proteins	Workman & Weyer, 2008
6844	NH (Secondary Amine)	Secondary Amines	Workman & Weyer, 2008
6849	N-H Str. first overtone	CONH2	NIRCal 5.4
6849	Hydroperoxide	Standards	Holman & Edmondson, 1956
6849	Hydroperoxide	Fatty Acid Standards	Holman et al., 1958
6850	Intramolecularly Bonded Alcohol OH	Alcohols	Workman & Weyer, 2008
6887	Primary Alcohol OH	Primary Alcohols	Workman & Weyer, 2008
6897	Moisture	Marine & Linseed Oil Dietary Supplements	Berzaghi et al., 2005
6897	Water	Ground Beef from Grass or Grain Fed Cattle	Realini et al., 2004
6897	O-H Str. first overtone	Starch,H2O	NIRCal 5.4
6897	Carbonyl	Ketones and Aldehydes	Workman & Weyer, 2008
6897	Polymeric OH	Starch / Polymeric Alcohols	Workman & Weyer, 2008

Wavenumber (cm^{-1})	Feature	Compound or Lipid System	Reference
6916	2x C-H Str.+ C-H def.	Aromatic	NIRCal 5.4
6935	Acid Value Calibration Wavelength	Soybean Oil	Cho et al., 1998
6940	Sugar OH	Crystalline Sucrose	Workman & Weyer, 2008
6944	2x C-H Str.+ C-H def.	CH	NIRCal 5.4
6944	Methylene CH	Aliphatic Compounds	Workman & Weyer, 2008
6944	OH	Lean Fish	Karlsdottir et al., 2014
6950	Peroxides (increase)	Olive; Sunflower; Rapeseed	Wojcicki et al., 2015
6964	Polar Content Determination of oil Heated at 170-180°C	Sunflower Oil (Fried)	El-Rafey et al., 1988
6964	Linoleic Acid Content Determination of oil Heated at 200-210°C	Sunflower Oil (Fried)	El-Rafey et al., 1988
6964	Linoleic Acid Content Determination of oil Heated at 200-210°C	Sunflower Oil (Fried)	El-Rafey et al., 1988
6969	OH	Palm; Rapeseed	Kaddour et al., 2006
6974	Polymer Content Determination of oil Heated at 170-180°C	Sunflower Oil (Fried)	El-Rafey et al., 1988
6993	O-H Str. first overtone	Sucrose, Starch	NIRCal 5.4
6995	NH (Primary Amides)	Primary Amides	Workman & Weyer, 2008
7018	OH	Pork Loin	Park et al., 2008
7042	O-H Str. first overtone	ArOH	NIRCal 5.4
7042	Alcohol Hydroxyl (intense)	Standards	Holman & Edmondson, 1956
7042	Hydroxyl	Fatty Acid Standards	Holman et al., 1958
7057	2x C-H Str.+ C-H def.	Aromatic	NIRCal 5.4
7060	Alcohol OH (hydrogen bonded)	Alcohols	Workman & Weyer, 2008

Wavenumber (cm^{-1})	Feature	Compound or Lipid System	Reference
7062	Polar Content Determination of oil Heated at 200-210°C	Sunflower Oil (Fried)	El-Rafey et al., 1988
7062	Polar Content Determination of oil Heated at 200-210°C	Sunflower Oil (Fried)	El-Rafey et al., 1988
7065	Alcohol OH (non-bonded)	Alcohols	Workman & Weyer, 2008
7067	2x C-H Str.+ C-H def.	CH2	NIRCal 5.4
7067	Methylene CH	Aliphatic Compounds	Workman & Weyer, 2008
7068	Hydroperoxides	Olive; Sunflower; Rapeseed	Wojcicki et al., 2015
7082	Polar Content Determination of oil Heated at 200-210°C	Sunflower Oil (Fried)	El-Rafey et al., 1988
7085	Methylene CH	Aliphatic Compounds (Branched)	Workman & Weyer, 2008
7090	OH (non-bonded)		Workman & Weyer, 2008
7092	O-H Str. first overtone	ROH	NIRCal 5.4
7092	Methylene CH	Aliphatic Compounds (Linear)	Workman & Weyer, 2008
7092	Alcohol OH	Alcohols	Workman & Weyer, 2008
7143	OH	Fish Oil	Cozzolino et al., 2005
7163	Methyl CH	Aliphatic Compounds (Branched)	Workman & Weyer, 2008
7163	Polymer Content Determination of oil Heated at 200-210°C	Sunflower Oil (Fried)	El-Rafey et al., 1988
7168	2x C-H Str.+ C-H def.	CH2	NIRCal 5.4
7168	Methylene CH	Aliphatic Compounds	Workman & Weyer, 2008
7194	Methyl CH	Aliphatic Compounds (Linear)	Workman & Weyer, 2008
7267	Polar Content Determination of oil Heated at 170-180°C	Sunflower Oil (Fried)	El-Rafey et al., 1988
7353	CH	Palm; Rapeseed	Kaddour et al., 2006
7353	2x C-H Str.+ C-H def.	CH3	NIRCal 5.4
7353	Methyl CH Combination	Aliphatic Compounds	Workman & Weyer, 2008

Wavenumber (cm^{-1})	Feature	Compound or Lipid System	Reference
8143	Polymer Content Determination of oil Heated at 170-180°C	Sunflower Oil (Fried)	El-Rafey et al., 1988
8163	C-H Str. Second overtone	CH	NIRCal 5.4
8163	CH (Secondary or Tertiary Carbon)	Aliphatic Compounds	Workman & Weyer, 2008
8197	Polymer Content Determination of oil Heated at 170-180°C	Sunflower Oil (Fried)	El-Rafey et al., 1988
8230	C-H Str. Second overtone	CH2	NIRCal 5.4
8230	Methylene CH	Aliphatic Compounds	Workman & Weyer, 2008
8258	Methylene CH	Aliphatic Compounds	Workman & Weyer, 2008
8264	CH	Palm; Rapeseed	Kaddour et al., 2006
8278	CH	Fish Oil	Cozzolino et al., 2005
8285	CH	Olive; Maize; Seed; Sunflower	Armenta et al., 2007
8333	CH2	Extra Virgin Olive Oil	Manley & Eberle, 2006
8333	CH	Pork Loin	Park et al., 2008
8368	C-H Str. Second overtone	CH3	NIRCal 5.4
8368	Methyl CH	Aliphatic Compounds	Workman & Weyer, 2008
8375	Methyl CH	Aliphatic Compounds	Workman & Weyer, 2008
8425	Protein	Ground Beef from Grass or Grain Fed Cattle	Realini et al., 2004
8547	C-H Str. Second overtone	HC=CH	NIRCal 5.4
8547	Alkene CH (CH=CH)	Alkenes	Workman & Weyer, 2008
8591-8651	Methyl Asymmetric CH Stretch		Osborne & Fearn, 1986
8600	CH2 (decrease)	Olive; Sunflower; Rapeseed	Wojcicki et al., 2015
8621	CH	Lean Fish	Karlsdottir et al., 2014
8621	Carbonyl	Aliphatic Compounds	Workman & Weyer, 2008
8681	C-H Str. Second overtone	CH3	NIRCal 5.4

Wavenumber (cm^{-1})	Feature	Compound or Lipid System	Reference
8730-8850	Alkene 2nd Overtone		Siesler, et al., 2002
8757-8850	Alkene 2nd Overtone		Osborne & Fearn, 1986
8930-9660	H-bonded Secondary Amide	Proteins	Siesler, et al., 2002
8937-9025	H-bonded Secondary Amide	Proteins	Osborne & Fearn, 1986
9116	2x C-H Str.+2x C-C Str.	Cyclopropane	NIRCal 5.4
9259	2x C-H Str.+2x C-C Str.	Benzene	NIRCal 5.4
9386	OH Combination	Water or Alcohols	Workman & Weyer, 2008
9434	N-H Str. Second overtone	RNH2	NIRCal 5.4
9497	2x C-H Str.+2x C-H def.+(CH2)n	CH2	NIRCal 5.4
9550	OH (hydrogen bonded)	Alcohols	Workman & Weyer, 2008
9600	Methylene CH	Aliphatic Compounds (Branched)	Workman & Weyer, 2008
9606	Methylene CH	Aliphatic Compounds (Linear)	Workman & Weyer, 2008
9643	2x C-H Str.+2x C-H def.+(CH2)n	Oil	NIRCal 5.4
9709	N-H Str. Second overtone	RNH2	NIRCal 5.4
9720	OH and CO Combination in Polyfunctional Alcohols	Ethers and Esters containing Alcohol Group	Workman & Weyer, 2008
9794	Methyl CH	Aliphatic Compounds (Branched)	Workman & Weyer, 2008
9795	Methyl CH	Aliphatic Compounds (Linear)	Workman & Weyer, 2008
9804	2x N-H Str.+2x amide	Protein	NIRCal 5.4
9852	2x C-H Str.+3x C-H def.	CH3	NIRCal 5.4
9940	Tertiary Alcohol OH		Workman & Weyer, 2008
9960	Secondary Alcohol OH		Workman & Weyer, 2008
10000	OH	Pork Loin	Park et al., 2008
10000	O-H Str. Second overtone	ArOH	NIRCal 5.4
10040	Primary Alcohol OH		Workman & Weyer, 2008
10417	CH	Lean Fish	Karlsdottir et al., 2014

APPENDIX B

Statistical Measures of Model Fitness

Throughout the process of calibration and validation, a variety of statistical measures are applied to determine the best models based on the calibration samples and how well each model fits the reference assay data for additional samples. During validation the predictions from the NIR model ($\hat{y}$) and the reference assay data (y) represent dependent and independent variables, respectively (Williams, 2001). The statistical measures of fit encompass both the accuracy and precision of the model. It is critical to interpret them correctly to avoid gross errors in the predictive ability of the model.

B.1 Defining Error to Assess Model Fitness

The procedure to create an NIR model is an inherently comparative process which relies upon the collection of valid data for the quantity of interest using an accurate reference assay. Assuming a set of ten samples are measured using a certain reference assay, an NIR model would ideally provide the same values and a plot of results from NIR versus the reference assay would fall on a line with a slope of 1 and zero intercept. Although such a plot in the case where reference assay values are 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10 would ideally look like Figure B.1a, in reality the results predicted from the NIR model will deviate from the ideal relationship, yielding a set of residuals (Figure B.1b).

The attributes of the distribution of residuals indicate the fitness of the NIR model in quantifying the analyte of interest. Distributions of the original data and the residuals from the example in Figure B.1 are shown in Figure B.2. The mean of the residuals is known as the bias:

$$BIAS = \sum_{i=1}^{N} (\hat{y}_i - y_i) \,/N$$

The bias is a measure of the accuracy of the NIR model, and its ideal value is zero (NIRCal 5.4 Software Manual). The standard deviation of the residuals is generically known as the standard

error of the estimate (SEE) (Naes et al., 2002). The SEE is more specifically termed the standard error of calibration (SEC) when applied to calibration samples, the standard error of performance (SEP) or standard error of validation (SEV) when applied to validation samples, and the standard error of cross validation (SECV) when samples are used for validation in a cross-validation procedure (Naes et al., 2002; NIRCal 5.4 Software Manual).

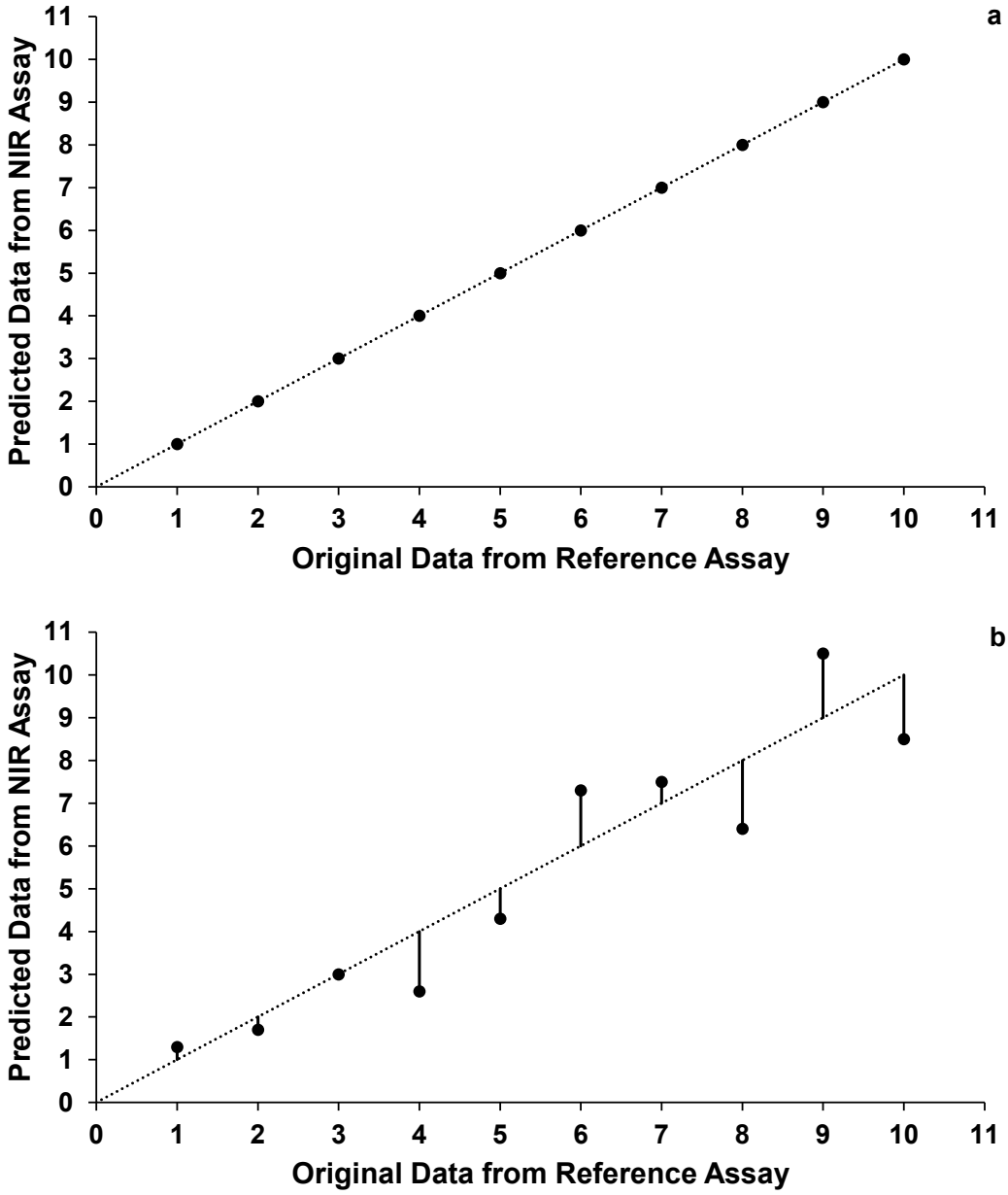


Figure B.1. Correlation of NIR predictions with reference assay values: (a) Ideal case and (b) Real case.

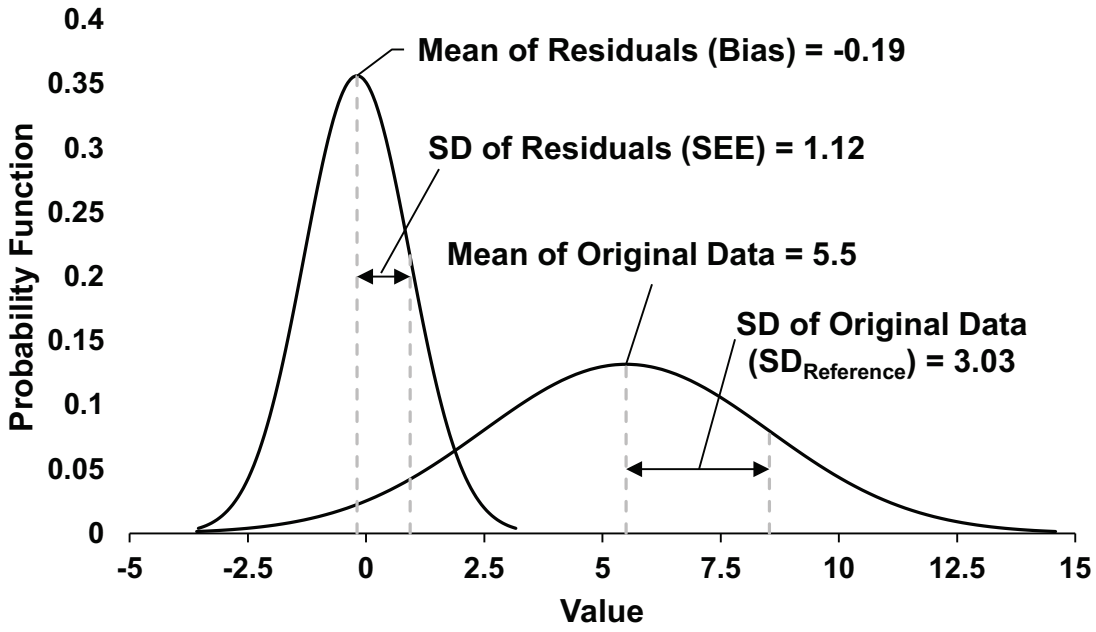


Figure B.2. Distributions of original data and residuals from plot of NIR predicted values vs reference assay values (Figure B.1b).

The ratio of the standard deviation of the reference assay data to SEE for the validation samples (whether SEP, SEV or SECV) is known as the residual predictive deviation (RPD) (Williams, 2001):

$$RPD = \frac{SD_{Reference}}{SEE}$$

As the SEE decreases relative to the standard deviation of the reference data, the predictive ability of the model improves and the RPD rises. A general interpretation of the value of RPD versus predictive ability is shown in Table B.1. (Williams, 2001). The RPD for the example distribution data in Figure B.2 is (3.03 / 1.12) = 2.71, indicating that such a model is suitable only for very rough screening (Table B.1.).

B.2 Coefficients of Correlation (r) and Determination (r^2)

A plot of the predicted ($\hat{y}$) versus reference assay values (y) for the samples ideally provides a line with a slope of 1.0 and a zero intercept (Figure B.1a). The correlation coefficient

Table B.1 Interpretation of residual predictive deviation (RPD) [From Williams, 2001].

RPD Value	Classification	Application
0.0 – 2.3	Very Poor	Not Recommended
2.4 – 3.0	Poor	Very Rough Screening
3.1 – 4.9	Fair	Screening
5.0 – 6.4	Good	Quality Control
6.5 – 8.0	Very Good	Process Control
8.1+	Excellent	Any Application

(r) indicates the degree of linear relationship between ($\hat{y}$) and (y) and should be as close to 1.0 as possible (Naes et al., 2002; Williams, 2001). The coefficient of determination (r^2) indicates the proportion of variance in ($\hat{y}$) explained by (y) and should be as close to 1.0 (100%) as possible (Williams, 2001). In reality, errors in both the reference and NIR data sets affect each of these measures (Naes et al., 2002).

B.3 Q-Value

The NIRCal chemometric software package from Buchi uses a penalty scoring index known as the Q-Value to assess the fit of both qualitative and quantitative models to calibration and validation data sets (NIRCal 5.4 Software Manual). The general formula for the Q-Value is:

$$Q - Value = \frac{1}{(1 + \sum wv)}$$

Here, w indicates the weight ascribed to a penalty while v is the number of samples subject to that penalty for qualitative models or the value of the parameter corresponding to that penalty calculated from the data for quantitative models. The Q-value ranges from 0 to 1, with higher values indicating better models. Buchi advises that a useful model typically has a Q-Value of at least 0.8. The penalties and weights for assessing qualitative and quantitative models are shown in Tables B.2 and B.3, respectively.

Table B.2 Q-Value criteria for qualitative models [From NIRCal 5.4 Software Manual].

Penalty	Ideal Value	Weight
C-Set False Identified (Calibration Sample in Wrong Cluster)	0	10
C-Set Not Identified (Calibration Sample Outside All Clusters)	0	10
V-Set False Identified (Validation Sample in Wrong Cluster)	0	5
V-Set False Identified (Validation Sample Outside All Clusters)	0	1
Cluster Index (Samples of Same Type Should be in Single Cluster)	0	1
Property Uniformity (Even Spread of Samples Within Clusters)	Small	1
Property Interference (Independence of Clusters from Each Other)	Small	0.1

Table B.3 Q-Value criteria for quantitative models [From NIRCal 5.4 Software Manual].

Penalty	Formula	Ideal Value	Weight
Rejection of Known	Number of C-Set Spectra with Residual too Large	0	10
Rejection of Unknown	Number of V-Set Spectra with Residual too Large	0	1
Relative Consistency	$\frac{\lvert(SEC - SEP)\rvert}{(\lvert SEP\rvert + 1)}$	C-Set and V-Set Have Similar Low Standard Error Terms	2
Weighted BIAS	$\frac{\lvert V - Set\ BIAS\rvert}{\lvert Range\rvert}$	Absolute V-Set BIAS Low	2
Validity	1 – (V-Set Regression Coefficient)	V-Set Regression Coefficient Near 1	1
Comparability	|(C-Set Regression Coefficient) – (V-Set Regression Coefficient)|	C-Set and V-Set Have Similar High Regression Terms	1
Precision	$\frac{SEP}{\lvert Range\rvert}$	Standard Error of the V-Set Low	1
Weighted Accuracy	$\sqrt{\frac{\left(\frac{\lvert V - Set\ RSS\rvert}{\lvert Range\rvert}\right)}{(\#\ of\ V - Set\ Samples)}}$	V-Set Residual Sum of Squared Error Low	1