

Development of Near  
Infrared Spectroscopy  
Methods to Assess Key  
Sultana Quality  
Parameters at the  
Processing Line

————— O —————

Chris Collins

Doctor of Philosophy

VICTORIA UNIVERSITY OF TECHNOLOGY



3 0001 00189 8057

2008  
VU

# **Development of Near Infrared Spectroscopy Methods to Assess Key Sultana Quality Parameters at the Processing Line**

**A thesis submitted for the degree of Doctor of Philosophy**

**Chris Collins**

**B.Sc. (Chemical Science, Deakin University) Honours**



**School of Molecular Sciences**

**Victoria University**

**Werribee**

**Victoria Australia**

**December 2006**

## Declaration

I, Christopher Collins, declare that the PhD thesis entitled "*Development of Near Infrared Spectroscopy Methods to Assess Key Sultana Quality Parameters at the Processing Line*" is no more than 100,000 words in length, exclusive of tables, figures, appendices, references and footnotes. The thesis contains no material that has been submitted previously, in whole or in part, for the award of any other academic degree or diploma. Except where otherwise indicated, this thesis is my own work.

Signature:

Date: 15/1/09

## Abstract

The Australian Sultana industry uses a visual assessment to classify sultanas, based on overall colour and relative uniformity of colour within a representative sample of 100 berries. A technique that gives a more precise indication of fruit quality is required by the industry. A key requirement is the rapid assessment of fruit after processing, to enable the processor to more readily meet the requirements of the buyer. The intention is to also meet the increasing demand of consumers who are becoming increasingly discerning and regulators who require more detailed nutritional information provided on food packaging.

Samples from the 2001 to 2004 seasons of processed sultanas, provided by the packing companies, were scanned from 400 to 2500nm using a Foss NIRSystems 6500 Spectrometer, then analysed by a series of physicochemical techniques: CIE tricolour stimulus values, dew point, titratable acidity and crude protein by Kjeldahl's procedure. These reference values were used to develop calibrations utilising Multiple Linear Regression (MLR) and Partial Least Squares (PLS) regression analysis techniques. The resultant calibrations demonstrated useful performance, for the rapid assessment of the quality parameters of interest.

The final PLS calibration for CIE  $b^*$ , the red-to-green colour axis of the CIE system important in determining maturity of the picked fruit, an important during grading of the fruit, achieved an  $R^2$  of 0.787, with a validation set  $R^2$  of 0.789 (SEP=0.385, Bias= 0.366). The PLS calibration for nitrogen by Kjeldahl's procedure attained an  $R^2$  of 0.859, with a validation set  $R^2$  of 0.758 (SEP=0.315, Bias= 0.0226). In the past these parameters proved challenging to develop calibrations for, as there was often little variation within a sample set, and in the case of Nitrogen, the analyte was present in small quantities.

The industry also requires a means of assessing fruit maturity or "full bodied ness" of unprocessed fruit. Maturity is currently used to grade sun dried

natural sultanas (called “raisins” in the United States of America), as fruit dried in this manner develop a uniform blue-black colour that cannot be distinguished visually. Natural sultanas have been graded using an airstream sorter, which separate the fruit sample into fractions of different densities by adjusting the flow rate of the air in the sorter. Airstream sorting has established relationships with titratable acidity and CIE parameters.

In this experiment, air-steam sorting was used in conjunction with harvesting fruit representatively over seven weeks, at weekly intervals, from an experimental plot at CSIRO Merbein, to provide a diverse range of fruit maturities for this calibration. Full spectra (400-2500nm) were collected in diffuse reflectance mode (coarse sample cell) using a NIRSystems 6500 Near Infrared Spectrophotometer, then CIE tricolour stimulus, dew point and titratable acidity were obtained. The spectra were pre-treated using N-point smooth and 2<sup>nd</sup> derivative, calibrations were developed using MLR and PLS in Vision (v.2.22 Foss NIRS Proprietary Software) as were the above calibrations. The resultant calibrations demonstrated potential as a means of fruit maturity grading of unprocessed sultanas.

The PLS calibration for Titratable acidity of unprocessed airstream sorted samples attained an  $R^2$  of 0.862, with a validation set  $R^2$  of 0.822 (SEP=35.0, Bias= -3.84). The attainment of robust NIR calibrations for the parameters of interest in whole sultana samples will benefit the Dried Vine Fruit industry greatly, and is a significant contribution to the body of knowledge and application of diffuse reflectance near infrared spectroscopy in the rapid assessment of agricultural products. In all the constituents in processed fruit it is the first set of NIRS calibrations successfully developed, and in the case of titratable acidity it is the first time a robust calibration has been developed in either processed or unprocessed fruit.

## Acknowledgements

Throughout my project a great many people have contributed and I now take this opportunity to thank them.

I'd like to thank my principal supervisor Assoc. Prof Mary Millikan of Victoria University, and my co-supervisor Mr Peter Clingeffer of CSIRO, Merbein for their guidance and support.

I would like to thank Dried Fruit Research and Development Council (DFRDC), Horticulture Australia Limited, (HAL) formerly Rural Industries Research and Development Council (RIRDC) and Victoria University (VUPSS Grant) for their financial support. Thanks are due to Mr. Ross Skinner, Executive Director (HAL), Mildura for his continued industry liaison.

Mr. David Swain, Sunbeam Foods, for support as our industry partner and supplying our samples.

Mrs Debra Thompson from the DFRDC Quality Centre for samples and use of the laboratory for industry demonstrations.

Ms Caroline Tarr, Mr David Emmanuelli and Ms Karen Connolly for growing, harvesting, drying and sorting the samples for the airstream-sorting chapter.

Mr Paul Brimmer, Mrs Jackie Brimmer (*nee* Boschenok) and Mr Jim Jasper of Foss Pacific for NIRS technical support and their advice.

Mr Lou Culpepper (formerly of Linbrook, the instrument supplier), Murray Goulburn Cooperative Company Ltd for advice regarding Vision (NIRSystems) Software.

Ms Ayla Boz for "showing me the ropes" early in my project.

The technical staff at VU, in particular Mr Joe Pelle, for their patience and support.

Dr Patrick Wright, formerly of Melbourne University, for his words of wisdom and his gentle patriarchal manner during my time at WT Kendall Hall.

Mr Craig Kyngdon of Melbourne University, a good mate whose infectious positive nature saw me through moments of self-doubt.

Mr Simon May, the ultimate opinionated sounding board

My mother and father for their constant support and belief in me.

All my kin, friends and peers whose help and friendship has guided me through my life.

## **Contents**

<b>Declaration</b>	<b>I</b>
<b>Abstract</b>	<b>II</b>
<b>Acknowledgements</b>	<b>IV</b>
<b>Contents</b>	<b>V</b>
<b>List of tables</b>	<b>XIII</b>
<b>List of diagrams</b>	<b>XXI</b>
<b>List of Abbreviations</b>	<b>XXVII</b>
<b>1.0 Introduction</b>	<b>1</b>
1.1 Contribution to the body of knowledge.	1
1.2 Research problem and hypothesis questions	3
1.3 Justification of the research.	3
1.4 Outline of the thesis.	5
1.4.1 Chapter 1: introduction	5
1.4.2 Chapter 2: literature review	5
1.4.3 Chapter 3: methodology	5
1.4.4 Chapter 4: Results and Discussion I: Quality Assurance of Processed Sultanas – CIE L*, a* and b* Values	6
1.4.5 Chapter 5: Results and Discussion II: Quality Assurance of Processed Sultanas - Water Activity, Protein, Titratable Acidity and Lipids.	6
1.4.6 Chapter 6: Results and discussion (III): Quality Assurance of Unprocessed Sultanas - Airstream Sorting	6
1.4.7 Chapter 7 General Discussion and Overall Conclusions	6
1.5 Bibliography	7

<b>2.0</b>	<b>Literature Review</b>	<b>8</b>
2.1	Origin of sultana production	8
2.2	The drying process	9
	2.2.1 Sun drying	9
	2.2.2 Rack drying	9
	2.2.3 Trellis Drying	10
	2.2.4 Chemical treatment	10
2.3	Processing	12
2.4	Grading	13
2.5	Factors affecting quality	14
	2.5.1 Enzymatic Browning	14
	2.5.2 Sunlight Induced Browning	14
	2.5.3 Cultivar	15
	2.5.4 Moisture Content	15
	2.5.5 Storage Time	15
2.6	Analytical Procedures	16
	2.6.1 Kjeldahl's Procedure, Crude Protein	16
	2.6.2 Dew Point	16
	2.6.3 Chromameter	17
	2.6.4 Airstream sorting technique and grading	17
	2.6.5 Airstream sorting and maturity	18
2.7	Near infrared spectrophotometry	19
	2.7.1 Molecular spectroscopy theory	19
	2.7.2 Measurement and vibrational modes	19
	2.7.3 Near Infrared theory	20
	2.7.4 Applications of near-infrared spectroscopy	21
2.8	Chemometrics	24
	2.8.1 Multiple Linear Regression	24
	2.8.2 Industrial and agricultural applications of MLR calibrations	24
	2.8.3 Partial Least Squares	26
	2.8.4 Industrial and agricultural applications of PLS calibrations	28

2.9	Applications of NIRS	30
2.9.1	Colour and NIRS	30
2.9.2	Dew Point and NIRS	33
2.9.3	Lipids and NIRS	34
2.9.4	Titrateable acidity and NIRS	34
2.9.5	Protein and NIRS	35
2.10	Aims	41
2.10.1	General aims of the project	41
2.10.2	Specific aims of the project	41
2.11	Bibliography	42
<b>3.0</b>	<b>Materials and Methods</b>	<b>56</b>
3.1	Introduction	56
3.2	Part I: NIRS assessment of processed sultanas	57
3.2.1	Sampling Procedure	57
3.2.2	Sample Preparation	57
3.2.3	Specific preparation for Titrateable acidity experiment	58
3.2.4	Specific preparation for Kjeldahl's procedure	58
3.2.5	Specific preparation for total lipids determination	58
3.3	Experimental Analysis	59
3.3.1	CIE tricolour stimulus	59
3.3.2	Water activity	59
3.3.3	Titrateable Acidity	59
3.3.4	Total Nitrogen	60
3.3.5	Total Lipids determination	62
3.3.6	NIRS Analysis	63
3.4	Part II: NIRS assessment of the maturity of unprocessed fruit	64
3.4.1	Introduction	64
3.5	Sampling Procedure	65
3.6	Sample Preparation	66
3.6.1	General technique	66
3.6.2	Sample Preparation for destructive analytical	

techniques	66
3.6.3 Airstream Sorting	66
3.7 Experimental analysis	67
3.7.1 CIE Tricolour stimulus	67
3.7.2 Water Activity by Dew Point	67
3.7.3 Titratable Acidity	67
3.7.4 NIRS Analysis	67
3.8 Calibration Development	68
3.9 Calibration optimisation and outlier removal	69
3.10 Bibliography	71
<b>4.0 Results and Discussion I: Quality Assurance of Processed Sultanas – CIE L* a* and b* Values</b>	<b>72</b>
4.1 Introduction	72
4.2 Objectives	73
4.3 Inter-correlation analysis	73
4.3.1 Year 1	73
4.3.2 Year 2	74
4.3.3 Year 3	75
4.3.4 Final combined calibration	76
4.4 Calibration development technique	77
4.5 Chemical assignments and NIRS	80
4.6 Tricolour Stimulus L*	83
4.6.1 Season one CIE L*	83
4.6.1.1 MLR calibration	83
4.6.1.2 PLS Calibration of CIE L*	87
4.6.2 Season two CIE L*	89
4.6.2.1 MLR Calibration	89
4.6.2.2 PLS Calibration	91
4.6.3 Season three CIE L*	93
4.6.3.1 MLR calibration	93
4.6.3.2 PLS Calibration	95
4.6.4 Combined Seasons calibrations of CIE L*	97

	4.6.4.1	MLR calibration	97
	4.6.4.2	PLS Calibration	101
4.7		Tricolour Stimulus a*	104
	4.7.1	Season one CIE a*	104
		4.7.1.1 MLR calibration	104
		4.7.1.2 PLS calibration	106
	4.7.2	Season two CIE a*	108
		4.7.2.1 PLS Calibration	108
		4.7.2.2 PLS Calibration	110
	4.7.3	Season three CIE a*	112
		4.7.3.1 MLR Calibration	112
		4.7.3.2 PLS Calibration	114
	4.7.4	Combined seasons CIE a*	116
		4.7.4.1 MLR Calibration	116
		4.7.4.2 PLS Calibration	120
4.8		Tricolour stimulus b*	123
	4.8.1	Season one CIE b*	123
		4.8.1.1 MLR	123
		4.8.1.2 PLS	125
	4.8.2	Season two CIE b*	127
		4.8.2.1 MLR Calibration	127
		4.8.2.2 PLS Calibration	129
	4.8.3	Season three CIE b*	131
		4.8.3.1 MLR calibration	131
		4.8.3.2 PLS Calibration	133
	4.8.4	Combined Seasons CIE b*	135
		4.8.4.1 MLR Calibration	135
		4.8.4.2 PLS Calibration	139
4.9		Comparison to literature examples	141
4.10		Conclusion	144
4.11		Bibliography	146

<b>5.0</b>	<b>Quality Assurance of Processed sultanas using NIRS (II)</b>	<b>149</b>
5.1	Introduction	149
5.2	Water activity	152
5.2.1	Season one water activity	152
5.2.1.1	MLR calibration	152
5.2.1.2	PLS calibration	154
5.2.2	Season two water activity	156
5.2.2.1	MLR Calibration	156
5.2.2.2	PLS calibration	157
5.2.3	Season three water activity	159
5.2.3.1	MLR calibration	159
5.2.3.2	PLS calibration	161
5.2.4	Combined Seasons Water Activity	163
5.2.4.1	MLR Calibration	163
5.2.4.2	PLS calibration	167
5.2.5	Comparison to cited literature	169
5.3	Kjeldahl protein	171
5.3.1	Season one Kjeldahl Protein	171
5.3.1.1	MLR Calibration	171
5.3.1.2	PLS calibration	173
5.3.2	Season two Kjeldahl protein	175
5.3.2.1	MLR calibration	175
5.3.2.2	PLS Calibration	177
5.3.3	Season three Kjeldahl protein	179
5.3.3.1	MLR Calibration	179
5.3.3.2	PLS calibration	181
5.3.4	Combined seasons Kjeldahl protein	183
5.3.4.1	MLR calibration	183
5.3.4.2	PLS Calibration	187
5.3.5	Comparison to cited literature	189
5.4	Titatable Acidity	193
5.4.1	Season one titratable acidity	193
5.4.1.1	MLR Calibration	193
5.4.1.2	PLS Calibration	195

5.4.2	Season two titratable acidity	197
5.4.2.1	MLR calibration	197
5.4.2.2	PLS calibration	199
5.4.3	Season three titratable Acidity	201
5.4.3.1	MLR calibration	201
5.4.3.2	PLS Calibration	203
5.4.4	Combined seasons titratable acidity	205
5.4.4.1	MLR Calibration	205
5.4.4.2	PLS calibration	209
5.4.5	Comparison to cited literature	211
5.5	Total lipids	213
5.5.1	First season Total Lipids	213
5.5.1.1	MLR Calibration	213
5.5.1.2	PLS Calibration	216
5.5.2	Comparison to cited literature	217
5.6	Initial Conclusions	219
5.7	Bibliography	221
<b>6.0</b>	<b>Results and discussion (III): Airstream Sorting</b>	<b>225</b>
6.1	Introduction	225
6.2	Preliminary interpretation of data	226
6.2.1	Airstream sorting: trends within parameters	226
6.2.2	Proportion of fruit versus Harvest time	228
6.2.3	Colour versus Harvest time	230
6.2.4	Water activity	233
6.2.5	Titratable acidity	234
6.2.6	Titratable acidity calibration development	235
6.3	Airstream-sorting NIRS calibration development	236
6.3.1	CIE L*	236
6.3.1.1	MLR calibration of CIE L*	236
6.3.1.2	PLS calibration of CIE L*	238

6.3.2	CIE a*	240
6.3.2.1	MLR Calibration of CIE a*	240
6.3.3	CIE b*	242
6.3.3.1	MLR calibration of CIE b*	242
6.3.3.2	PLS calibration of CIE b*	244
6.3.4	Water activity	246
6.3.4.1	MLR calibration of Aw	246
6.3.5	Titrateable Acidity	248
6.3.5.1	MLR calibration of TA	248
6.3.5.2	PLS calibration of TA	252
6.3.5.3	Comparison to literature examples	254
6.4	Conclusion	255
6.4.1	Harvest time versus fruit maturity	255
6.4.2	Proportion of Fruit Versus Harvest time	255
6.4.3	Colour versus Harvest time	256
6.5	Bibliography	257
<b>7.0</b>	<b>General Discussion and Overall Conclusions</b>	<b>260</b>
7.1	Assessment of processed sultana quality using NIRS	260
7.2	Maturity assessment of unprocessed sultanas using NIRS	264
7.3	Theoretical implications	265
7.4	Practical implications	266
7.5	Future research	267
7.6	Bibliography	269
<b>Appendix 1</b>	<b>Season 1 laboratory data</b>	<b>271</b>
<b>Appendix 2</b>	<b>Season 2 laboratory data</b>	<b>274</b>
<b>Appendix 3</b>	<b>Season 3 laboratory data</b>	<b>276</b>
<b>Appendix 4</b>	<b>laboratory data for airstream sorting</b>	<b>280</b>
<b>Appendix 5:</b>	<b>Experimental procedures undertaken at CSIRO, Merbein for sultana maturity by air-steam sorting</b>	<b>283</b>

## List of Tables

Table 4.1	Inter-correlations between constituents of the first seasons' laboratory data	73
Table 4.2	Matrix of correlations between constituents of the second seasons' laboratory data	74
Table 4.3	Matrix of correlations between constituents of the third seasons' laboratory data	75
Table 4.4	Matrix of correlations between constituents of the combined seasons' laboratory data	76
Table 4.5	Laboratory parameters summary of season one CIE L*	83
Table 4.6	The wavelengths of the first seasons MLR calibration of CIE L* (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	85
Table 4.7	The factors of the first seasons PLS calibration (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	87
Table 4.8	Laboratory parameters summary of season two L*	89
Table 4.9	The wavelengths contributing to the second seasons MLR calibration (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	90
Table 4.10	The factors contributing to the second seasons PLS calibration of CIE L* (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	91
Table 4.11	Laboratory parameters summary of season three L*	93
Table 4.12	The wavelengths contributing to the third seasons MLR CIE L* calibration (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	94
Table 4.13	The factors contributing to the third seasons PLS calibration of CIE L* (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	96
Table 4.14	Laboratory parameters summary of combined L*	97

Table 4.15	Wavelengths contributing to the combined seasons MLR calibration of CIE L* (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	99
Table 4.16	Factors contributing to the combined seasons PLS calibration of CIE L* (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	102
Table 4.17	Laboratory parameters summary of season one a*	104
Table 4.18	Wavelengths contributing to the first seasons MLR calibration of CIE a* (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	105
Table 4.19	Factors contributing to the first seasons PLS calibration of CIE a* (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	106
Table 4.20	Laboratory parameters summary of season two a*	108
Table 4.21	Wavelengths used to contribute to the second seasons MLR calibration for CIE a* (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	109
Table 4.22	Factors contributing to second seasons PLS calibration of CIE a* (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	111
Table 4.23	Laboratory parameters summary of season three CIE a*	112
Table 4.24	Wavelengths contributing to the third seasons calibration MLR of CIE a* (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	113
Table 4.25	Factors contributing to the third seasons PLS calibration for CIE a* (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	114
Table 4.26	Laboratory parameters summary of combined a*	116
Table 4.27	Wavelengths contributing to the combined MLR calibration for CIE b* (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	118

Table 4.28	Factors contributing to the combined PLS calibration of CIE a* (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	121
Table 4.29	Laboratory parameters summary of season one b*	123
Table 4.30	Wavelengths contributing to the first seasons MLR calibration of CIE b* (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	124
Table 4.31	Terms used to optimise the first seasons PLS calibration of CIE b* (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	126
Table 4.32	Laboratory parameters summary of season two b*	127
Table 4.33	Wavelengths contributing to the second seasons MLR calibration of CIE b* (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	128
Table 4.34	Factors contributing to the second seasons PLS calibration of CIE b* (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	129
Table 4.35	Laboratory parameters summary of season three b*	131
Table 4.36	Wavelengths contributing to the third season MLR calibration of CIE b* (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	132
Table 4.37	Factors used to optimise the third season PLS calibration of CIE b* (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	133
Table 4.38	Laboratory parameters summary of combined b*	135
Table 4.39	Wavelengths contributing to the combined MLR calibration of CIE b* (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	137
Table 4.40	Factors contributing to the combined PLS calibration of CIE b* (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	140
Table 5.1	Laboratory parameters summary of season one Aw	152

Table 5.2	Wavelengths contributing to the MLR calibration of water activity (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	153
Table 5.3	Factors contributing to the PLS first seasons calibration of water activity (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	154
Table 5.4	Laboratory parameters summary of season two Aw	156
Table 5.5	Wavelengths contributing to the second seasons MLR calibration of water activity (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	157
Table 5.6	Factors contributing to the second seasons PLS calibration of water activity (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	158
Table 5.7	Laboratory parameters summary of season three Aw	159
Table 5.8	Wavelengths contributing to the third season calibration of water activity (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	160
Table 5.9	Factors contributing to the third seasons PLS calibration of water activity (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	162
Table 5.10	Laboratory parameters summary of combined Aw	163
Table 5.11	Wavelengths contributing to the combined seasons MLR calibration of water activity (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	166
Table 5.12	Factors contributing to the combined seasons PLS calibration of water activity (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	168
Table 5.13	Laboratory parameters summary of season one Kjeldahl protein	171
Table 5.14	Wavelengths contributing to the first seasons calibration of Kjeldahl protein (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	172

Table 5.15	Factors contributing to the first seasons PLS calibration of Kjeldahl protein (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	173
Table 5.16	Laboratory parameters summary of season two Kjeldahl Nitrogen	175
Table 5.17	Terms contributing to the second seasons' MLR calibration of Kjeldahl protein (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	176
Table 5.18	Factors contributing to the second seasons PLS calibration of Kjeldahl protein (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	177
Table 5.19	Laboratory parameters summary of season three Kjeldahl Nitrogen	179
Table 5.20	Wavelengths contributing to the third seasons MLR calibration of Kjeldahl protein (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	180
Table 5.21	Factors contributing to the third seasons PLS calibration of Kjeldahl protein (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	181
Table 5.22	Laboratory parameters summary of combined Kjeldahl protein	183
Table 5.23	Wavelengths contributing to the combined seasons MLR calibration of Kjeldahl protein (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	186
Table 5.24	Factors contributing to the combined seasons calibration of Kjeldahl protein (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	188
Table 5.25	Laboratory parameters summary of season one TA	193
Table 5.26	Wavelengths contributing to the first seasons MLR calibration for titratable acidity (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	194
Table 5.27	Factors contributing to the first seasons PLS calibration of titratable acidity (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	195

Table 5.28	Laboratory parameters summary of season two TA	197
Table 5.29	Wavelengths contributing to the second seasons MLR calibration of titratable acidity (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	198
Table 5.30	Factors contributing to the second seasons PLS calibration of titratable acidity (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	199
Table 5.31	Laboratory parameters summary of season three TA	201
Table 5.32	Wavelengths contributing to the third seasons MLR calibration of titratable acidity (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	202
Table 5.33	Factors contributing to the third seasons PLS calibration of titratable acidity (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	204
Table 5.34	Laboratory parameters summary of combined TA	205
Table 5.35	Wavelengths contributing to the combined MLR calibration of titratable acidity (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	208
Table 5.36	Factors contributing to the combined PLS calibration of titratable acidity (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	210
Table 5.37	Laboratory parameters summary of season one lipids	213
Table 5.38	Wavelengths contributing to the first seasons MLR percent lipids calibration (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	215
Table 5.39	Factors contributing to the first seasons' PLS percent lipids calibration (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	217
Table 6.1	Recovery proportion values of fruit at pressure settings for harvest dates.	227
Table 6.2	Recovered fruit fractions compared to date of harvest versus airstream sorting settings	229

Table 6.3	AS setting at harvest date versus average CIE L* value	230
Table 6.4	AS setting at harvest date versus average CIE a* value	231
Table 6.5	AS setting at harvest date versus average CIE b* value	232
Table 6.6	Harvest dates versus average dew point (water activity) values at airstream settings	233
Table 6.7	Harvest dates versus average TA values at airstream settings	234
Table 6.8	Laboratory parameters summary of CIE L*	236
Table 6.9	Wavelengths contributing to the airstream sorting MLR calibration of CIE L* (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	237
Table 6.10	Factors contributing to the airstream sorting PLS calibration of CIE L* (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	238
Table 6.11	Laboratory parameters summary of CIE a*	240
Table 6.12	Terms contributing to the airstream sorting MLR calibration of CIE a* (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	241
Table 6.13	Laboratory parameters summary of CIE b*	242
Table 6.14	Terms contributing to the airstream sorting MLR calibration for CIE b* (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	243
Table 6.15	Factors contributing to the airstream sorting PLS calibration of CIE b* (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	244
Table 6.16	Laboratory parameters summary of Aw	246
Table 6.17	Wavelengths contributing to the air steam sorting MLR calibration of water activity (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	247
Table 6.18	Laboratory parameters summary of Fruit Maturity TA	248

Table 6.19	Wavelengths contributing to the airstream sorting calibration of titratable acidity (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	251
Table 6.20	Factors making up the airstream sorting PLS calibration of titratable acidity (Pre-treatment: N-point smooth, 2 <sup>nd</sup> derivative)	253

## List of Diagrams

Figure 4.1	A residual plot illustrating outlier selection.	78
Figure 4.2	A plot of the first seasons calibration for CIE L* using MLR.	83
Figure 4.3	A plot of the first seasons calibration for CIE L* using PLS.	87
Figure 4.4	A plot of the second seasons calibration for CIE L* using MLR.	89
Figure 4.5	A plot of the second seasons calibration for CIE L* using PLS	91
Figure 4.6	A plot of the third seasons calibration for CIE L* using MLR	93
Figure 4.7	A plot of the combined seasons calibration for CIE L* using PLS.	95
Figure 4.8	A plot of the combined seasons calibration for CIE L* using MLR.	97
Figure 4.9	A plot of the combined seasons validation set for CIE L* using MLR.	97
Figure 4.10	Spectrum of the minimum (blue), maximum (red) and mean (yellow) samples of CIE L*	98
Figure 4.11	Close up of NIR spectra at the first wavelength of the combined MLR calibration	99
Figure 4.12	A plot of the combined seasons calibration for CIE L* using PLS	101
Figure 4.13	A plot of the combined seasons validation set for CIE L* using PLS	101
Figure 4.14	A plot of the first seasons calibration for CIE a* using MLR	104
Figure 4.15	A plot of the first seasons calibration for CIE a* using PLS	106
Figure 4.16	A plot of the second seasons calibration for CIE a* using MLR	108

Figure 4.17	A plot of the second seasons calibration for CIE $a^*$ using PLS	110
Figure 4.18	A plot of the third seasons calibration for CIE $a^*$ using MLR	112
Figure 4.19	A plot of the third seasons calibration for CIE $a^*$ using PLS	114
Figure 4.20	A plot of the combined seasons calibration for CIE $a^*$ using MLR	116
Figure 4.21	A plot of the combined seasons validation set for CIE $a^*$ using MLS	116
Figure 4.22	Spectra of the minimum (blue), maximum (red) and mean (yellow) of CIE $a^*$	117
Figure 4.23	Close up of NIR spectra at the first wavelength of the combined MLR calibration	117
Figure 4.24	A plot of the combined seasons calibration for CIE $a^*$ using PLS	120
Figure 4.25	The combined seasons validation set for CIE $a^*$ using PLS	120
Figure 4.26	A plot of the first seasons calibration for CIE $b^*$ using MLR	123
Figure 4.27	A plot of the first seasons calibration for CIE $b^*$ using PLS	125
Figure 4.28	A plot of the second seasons calibration for CIE $b^*$ using MLR	127
Figure 4.29	A plot of the second seasons calibration for CIE $b^*$ using PLS	129
Figure 4.30	A plot of the third seasons calibration for CIE $b^*$ using MLR	131
Figure 4.31	A plot of the third seasons calibration for CIE $b^*$ using PLS	133
Figure 4.32	A plot of the combined seasons calibration for CIE $b^*$ using MLR	135
Figure 4.33	A plot of the combined seasons validation for CIE $b^*$ using MLR	135

Figure 4.34	Spectra of the minimum (blue), maximum (red) and mean (yellow) samples of CIE $b^*$	136
Figure 4.35	Close up of NIR spectra at the first wavelength of the combined MLR calibration	137
Figure 4.36	A plot of the combined seasons calibration for CIE $b^*$ using PLS	139
Figure 4.37	The combined seasons validation set for CIE $b^*$ using PLS	139
Figure 5.1	A plot of the first seasons calibration of water activity using MLR	152
Figure 5.2	A plot of the first seasons calibration of water activity using PLS	154
Figure 5.3	A plot of the second seasons calibration for water activity using MLR	156
Figure 5.4	A plot of the second seasons calibration for water activity using PLS	157
Figure 5.5	A plot of the third seasons calibration for water activity using MLR	159
Figure 5.6	A plot of the second seasons calibration for water activity using PLS	161
Figure 5.7	A plot of the combined calibration for water activity using MLR	163
Figure 5.8	A plot of the combined validation set for water activity using MLR	163
Figure 5.9	Spectra of the minimum (blue), maximum (red) and mean (yellow) samples of water activity	164
Figure 5.10	Close up of NIR spectra at the first wavelength of the combined MLR calibration	165
Figure 5.11	A plot of the combined calibration for water activity using PLS	167
Figure 5.12	A plot of the combined validation set for water activity using PLS	167
Figure 5.13	A plot of the first seasons calibration for Kjeldahl protein using MLR	171

Figure 5.14	A plot of the first seasons calibration for Kjeldahl protein using PLS	173
Figure 5.15	A plot of the second seasons calibration for Kjeldahl protein using MLR	175
Figure 5.16	A plot of the second seasons calibration for Kjeldahl protein using PLS	177
Figure 5.17	A plot of the third seasons calibration for Kjeldahl protein using MLR	179
Figure 5.18	A plot of the third seasons calibration for Kjeldahl protein using PLS	181
Figure 5.19	A plot of the combined calibration of Kjeldahl protein using MLR	183
Figure 5.20	A plot of the combined validation set for protein using MLR	183
Figure 5.21	Spectra of the minimum (blue), maximum (red) and mean (yellow) samples of Kjeldahl protein	184
Figure 5.22	Close up of NIR spectra at the first wavelength of the combined MLR calibration	185
Figure 5.23	A plot of the combined calibration for Kjeldahl protein using PLS	187
Figure 5.24	A plot of the combined validation set for protein using PLS	187
Figure 5.25	A plot of the first seasons calibration for titratable acidity using MLR	193
Figure 5.26	A plot of the first seasons calibration for titratable acidity using PLS	195
Figure 5.27	A plot of the second seasons calibration for titratable acidity using MLR	197
Figure 5.28	A plot of the second seasons calibration for titratable acidity using PLS	199
Figure 5.29	A plot of the third seasons calibration for titratable acidity using MLR	201
Figure 5.30	A plot of the first seasons calibration for titratable acidity using PLS	203

Figure 5.31	A plot of the combined calibration for titratable acidity using MLR	205
Figure 5.32	The combined validation set for titratable acidity using MLR	205
Figure 5.33	Spectra of the minimum (blue), maximum (red) and mean (yellow) samples of titratable acidity	206
Figure 5.34	Close up of NIR spectra at the first wavelength of the combined MLR calibration.	207
Figure 5.35	A plot of the combined calibration for titratable acidity using PLS	209
Figure 5.36	The combined validation set for titratable acidity using PLS	209
Figure 5.37	A plot of the calibration for percent lipids using MLR	213
Figure 5.38	Spectra of the minimum (blue), maximum (red) and mean (yellow) samples of percent lipids	214
Figure 5.39	Close up of NIR spectra at the first wavelength of the combined MLR calibration	215
Figure 5.40	A plot of the calibration for percent lipids using PLS	216
Figure 6.1	A graph of fruit proportions recovered from samples harvested at different dates.	226
Figure 6.2	Fruit fractions recovered at an AS pressure setting of fruit harvested at different dates	228
Figure 6.3	Harvest dates versus average CIE L* values at airstream settings	230
Figure 6.4	Harvest dates versus average CIE a* values at airstream settings	231
Figure 6.5	Harvest dates versus average CIE b* values at airstream settings	232
Figure 6.6	Harvest dates versus dew point (water activity) values at airstream settings	233
Figure 6.7	Harvest dates versus titratable acidity (mL of 0.1M NaOH/100g sample) values at airstream settings	234
Figure 6.8	A plot of the MLR calibration for CIE L*	236
Figure 6.9	A plot of the PLS calibration for CIE L*	238

Figure 6.10	A plot of the MLR calibration for CIE a*	240
Figure 6.11	A plot of the MLR calibration for CIE b*	242
Figure 6.12	A plot of the PLS calibration for CIE b*	244
Figure 6.13	A plot of the MLR calibration of water activity	246
Figure 6.14	A plot of the calibration of Titratable acidity using MLR.	248
Figure 6.15	A plot of the validation set of titratable acidity using MLR.	248
Figure 6.16	Spectra of the minimum (blue), maximum (red) and mean (yellow) samples of titratable acidity	249
Figure 6.17	Close up of NIR spectra at the first wavelength of the combined MLR calibration	250
Figure 6.18	A plot of the calibration of titratable acidity using PLS	252
Figure 6.19	A plot of the validation set of titratable acidity using PLS	252

## List of Abbreviations

a*	CIE red-green coordinate
AOAC	Association of analytical chemists
AS	Airstream sorting
Aw	Water activity
b*	CIE blue-yellow coordinate
CIE	<i>Commission Internationale d'Eclairage</i>
FT-NIRS	Fourier transform near infrared spectroscopy
KP	Kjeldahl's procedure for protein determination
L*	CIE white-black coordinate
MCFT-NIRS	Multi channel Fourier transform near infrared spectroscopy
MIR	Mid infrared spectroscopy
MLR	Multi-linear regression
MPLS	Modified partial least squares
NIR	Near infrared
NIRS	Near infrared spectroscopy
NIT	Near infrared transmittance spectroscopy
PCA	Principal component analysis
PLS	Partial least squares
PPO	Polyphenol oxidase
R <sup>2</sup> (Cal)	Coefficient of determination of calibration set
R <sup>2</sup> (Val)	Coefficient of determination of validation set
SEC	Standard error of calibration
SECV	Standard error of cross validation
SEP	Standard error of prediction
SIMCA	Soft independent modelling of class analogy
TA	Titrateable acidity

## **1.0 Introduction**

### **1.1 Contribution to the body of knowledge.**

- The analyses of sultanas on the finishing line, prior to packaging, for a range of quality parameters has not been reported in the literature. In fact, colour grading by experienced personnel, after the fruit has passed over the metal detectors and Elbiscan laser scanners, is the only quality assessment measure that is currently undertaken. The Elbiscan laser scanner redirects objects that tumble past the laser scanner which have reflective properties that are outside preset parameters, this triggers an aimed jet of compressed air which send the foreign object into a waste container, allowing objects which have the reflective properties of sultanas to continue along the processing line.
  
- There are seven grades, know as Crown Grades by which the sultanas are classified [1]. The premium grade, Crown Grade Seven, represents 100% of golden berries and brings a premium price on the export market. The greater the percentage of dark berries present in the batch, the lower the Crown Grade. Industry personnel have given a high priority to the replacement of the subjective methods of sultana measurement by an unbiased instrumental method, which is one of the main aims of the project. In addition, consumers are becoming more discerning regarding the quality of food products while at the same time, regulatory authorities are requiring more nutritional information to be included on labels of processed foods.

- Method development will be required in this project, to adapt, extend and apply literature methods to the analyses for protein, lipids and sugar content, in sultanas containing the added parameter of dressing oil, that is coated on the finished product. Other parameters such as water activity and colour measurements will also be assessed. A very wide range of samples over three seasons will need to be analysed, to allow for seasonal differences that affect the quality of the fruit. The intellectual challenge will be to correlate the instrumental data with the NIR spectra, in the first instance, then to employ a range of different statistical methods to process the data. These methods include: Partial Least Squares (PLS), Multiple Linear Regression (MLR) or Principal Component Analysis (PCA), that will be applied to the spectra, to determine which of these approaches will give the most suitable correlation coefficient.

This statistical analysis is by no means a straight -forward process, as each individual parameter may require a different type of analysis. Once the calibration algorithms have been determined then the validation is undertaken. For example, samples that were not in the calibration are taken as the unknown set and the calibration equations developed are then used to predict the exactness of fit of the unknown data to the calibration curve. This work will be carried out over three years and then the calibrations for each single parameter will be obtained, by combining the data for each of the three years. Each calibration equation must be robust, taking into account the variations in each parameter due to climatic and seasonal effects.

The understanding of the correlation of the sultana quality values with the instrumental measurements, and in particular, to replace the subjective method of Crown Grade assessment will be an important contribution to the scientific literature and the Australian Dried Fruit Industry.

## **1.2 Research problem and hypothesis questions**

- *Is it feasible to use NIRS to assess processed fruit quality at the processing line?*
- *Is an NIR calibration of unprocessed fruit maturity based on airstream sorting possible?*

## **1.3 Justification of the research.**

- The Australian Dried Fruit Industry is a multi-million dollar export (2004-05: 6,626 tonnes worth \$15.4 million) industry and sultanas are exported to many countries including Italy, UK and Germany to name a few [2]. The main competitors to the Australian industry are Spain and Turkey where the Thompson Seedless grapes are also sprayed with a drying oil when on the vines, to assist in the rapid dehydration of the berries [3]. This is the procedure followed in Australia. It is important that Australia continues to improve the technology associated with the production and processing of light golden sultanas that bring premium prices on the world market.
- In order to reduce costs and streamline operations, it is necessary to adopt new technologies, as they become available. Particularly, if subjective assessment of fruit by trained personnel can be replaced by instrumental

methods of quality assessment. Current work at Victoria University that is nearing completion has addressed the assessment of sultanas at receipt for shelf life stability; this is necessary as they may be stored in the packing sheds for up to twelve months, unprocessed until they are sold [4,5].

- This project will, in contrast, to the previous one assess sultanas on the finishing-line, prior to packaging for a range of quality attributes. There is also the added complexity of modifying procedures to account for the dressing oil that is added to the fruit prior to packaging. One of the most important parameters is to replace the current Crown Grade Classification system, a subjective grading, by experienced personnel, with a Near Infrared assessment of sultana colour. Such an instrumental method of grading sultanas would assist the producers to reduce processing costs, as the instrumental method would be more rapid, unbiased and reproducible over a number of seasons.

## **1.4 Outline of the thesis.**

### **1.4.1 Chapter 1: introduction**

The introduction introduces the reader to the key research problems met by this project, outlines what approaches shall be taken to solve them and the detail of the techniques used.

### **1.4.2 Chapter 2: literature review**

The literature review gives the reader an introduction to the production of light coloured sultanas in Australia, an insight into the theory of molecular spectroscopy and Near Infrared spectroscopy in particular. The literature review examines the Laboratory techniques used in this project, and gives examples of these techniques used to assess a variety of parameters and applications. It also shows how near infrared spectroscopy can be used in conjunction with the statistical techniques of Chemometrics to analyse a variety of products for the parameters indicative of quality in Sultanas.

### **1.4.3 Chapter 3: methodology**

The first methodology chapter gives a detailed account of the sample preparation spectroscopic and analytical techniques used in the assessment of processed sultanas. This chapter also includes the process of developing a calibration using Vision (Version 2.22 Foss NIRSystems, Proprietary Software) from collecting spectra to calibration validation. This chapter also gives a detailed account of the techniques used to develop a means of assessing unprocessed sultanas for key maturity indicators.

#### **1.4.4 Chapter 4: Results and Discussion I: Quality Assurance of Processed Sultanas – CIE L\* a\* and b\* Values**

This chapter presents the results obtained for three seasons of CIE tricolour stimulus correlated to NIRS spectra.

#### **1.4.5 Chapter 5: Results and Discussion II: Quality Assurance of Processed Sultanas – Water Activity, Protein, Titratable Acidity and Lipids.**

This chapter presents the results obtained for three seasons  $A_w$ , titratable acidity, Kjeldahl protein and total lipids correlated to NIRS spectra.

#### **1.4.6 Chapter 6: Results and discussion (III): Quality Assurance of Unprocessed Sultanas - Airstream Sorting**

This chapter presents the results of the study to assess berry maturity of dried unprocessed fruit using NIRS in conjunction with chemometrics.

#### **1.4.7 Chapter 7: conclusions**

This chapter summarises the results of each study and indicates how these results have met the aims of this project. It also indicates how this will be implemented and future areas of study that may be required.

## 1.5 Bibliography

1. Grncarevic, M, Lewis, W. "External Colour of Dried Sultanas", *Food Technology in Australia* 1973 November: 562-565.
2. Australian Bureau of Statistics 2005: 1329.0
3. Uhlig, BA. "The Production of Light Coloured Dried Sultanas", *Die Weinwissenschaft: viticultural and enological sciences* 1996; 51(1): 37-39.
4. Frank, D. *Investigation of the biochemical basis of browning during the storage of sultanas, a PhD Thesis*. Victoria University, Werribee Campus; 2001.
5. Boz, A. *Multi-componet NIR analysis of dried vine fruit, an MSc Thesis*. Victoria University, Werribee Campus; 2002.

## **2.0 Literature Review**

### **2.1 Origin of sultana production**

The origin of sultana production is not well documented; sultanas are likely to have originated from Iran near a town called Souldanieh, the variety then spread across the Middle East into Asia and southern Russia [1]. The first introduction of sultanas into the Sunraysia region was during 1890's from the stock originating from vineyards established in the Cape of Good Hope, South Africa. The Sultana variety of grape proved well suited to the conditions of the Sunraysia region and became popular with growers as the vines grew well and the fruit dried quickly [1].

There are several regions within Australia that grow grapes for sultana production, however, the most significant are the Riverland and Sunraysia regions near the Victoria/New South Wales/South Australia border, where there is a climate of a low spring frost incidence, low spring rainfall and a hot, dry summer [1]. Despite its popularity, it was not until the end of the Great War that a significant industry developed, due to areas of the Riverland and Sunraysia regions being set aside as settlements for returned soldiers [1]. Total grape production for drying in Australia in the year 2004-05 was 135,412 tonnes, of which 6,626 tonnes went to export markets such as Canada, Germany, the United Kingdom and Italy [2].

## **2.2 The drying process**

### **2.2.1 Sun drying**

In many countries that produce sultanas, such as South Africa, drying methods have remained unchanged for centuries: the berries are placed on a surface that will store solar heat during the day and maintain a high temperature during the night, such as bare rock or concrete [3,4]. This process, while simple is slow and vulnerable to changes in weather and results in dark coloured sultanas less favoured by consumers.

### **2.2.2 Rack drying**

Sultana producing countries such as Australia have developed the use of drying racks, made of reinforced wire mesh suspended between posts, in some cases the racks are protected by a roof [5]. These drying racks evolved from the use of Hessian on the ground at the base of the vines, to dry the fruit following harvest [6]. This method has the advantage that multiple layers of drying racks can be supported on a drying frame, which allows a greater volume of fruit to be dried per area of ground allocated for this purpose [7]. It also has the advantage that if weather conditions change, it is possible to cover the fruit more quickly. In addition, this method is also less susceptible to contamination [8].

Following harvest the fruit are either bulk dipped or sprayed with emulsion before being placed on the drying racks for up to fourteen days [9]. Once dried the fruit are shaken and collected from the racks using a tractor-mounted mechanical shaker [5].

### **2.2.3 Trellis Drying**

Trellis drying or “summer pruning” allows the fruit to be harvested mechanically, reducing the inherent problems of employing a large number of casual pickers and allows varieties more sensitive to handling when fresh to be grown, like Merbein Seedless [1]. Contamination is reduced and enables large areas of fruit to be quickly dried and sent to the packing sheds, allowing growers to cope better with adverse weather conditions. The fruiting canes are cut, leaving behind next years’ replacement canes and no less than half the vines canopy. The fruit is then sprayed with drying emulsion prior to or immediately after the cane cutting. The dried fruit are ready to harvest when the fruit reaches 16% moisture, as indicated by the brittle stems. The fruit is then harvested mechanically but the remaining fruit missed by the harvester is picked by hand. Crown Fruit left on unpruned canes are either picked by hand prior to harvest, or pruned before they can develop. Pruning the crown fruit bunches does not affect the over all yield as it enhances the fruit remaining on the vine. The fruit is then dried to the desired moisture content on black plastic on the ground then sent for processing, as the fruit has a longer shelf life unprocessed [1].

### **2.2.4 Chemical treatment**

Chemical treatment of fruit has also been shown to increase drying times by making the skin of berries more permeable and results in berries of a bright golden colour, rather than the blue black fruit resulting from sun drying [1]. In Australia, chemical treatments of fruit with aqueous oil emulsion and

potassium carbonate has been widely used since 1969 [10], and results in fruit of a bright golden colour desirable to consumers without the use of harsh chemical treatment. The use of potassium carbonate aqueous emulsions originated in Mediterranean and continues to this day in Turkey, Iran and several other Middle-Eastern countries [10]. This is in contrast to the exposure of sultanas to sulphur dioxide in the United States, which produces consistently golden coloured sultanas. However, consumers are becoming increasingly aware of possible health risks caused by chemicals used in food processing [11].

In both the Trellis Drying and Hand Picking methods the Sultanas are treated with an emulsion treatment (called "Cold Dip"), which modifies the outer layer of lipid platelets, making it more permeable. This treatment reduces the drying time of fruit, from 4-5 weeks to 8-14 days, and has the additive affect of inhibiting the enzymatic browning mechanisms within sultanas that affect fruit quality, particularly colour [10,1]. Once the fruit has reached a moisture content of 18% they are ready for final drying to 13% then they are stored unprocessed until sold [12].

## **2.3 Processing**

After the fruit has been received by the packing sheds, it is first pre-cleaned, where the scalper riddle and push pull vacuum extractors remove cane fragments, fruit clumps and lightweight trash. Following this, the fruit is passed over magnets, which pick up any metal objects present. The fruit then passes through the cone, which removes cap stems. The cone consists of a rotating metal cone with protruding rub bars inside a stationary mesh cage, with a clearance between cone and cage of 2 to 2.5 cm, when the fruit pass through this clearance, the rotation of the cone removes the cap stems by abrasion.

The fruit passes through a riddle box, which removes any over or undersized objects from the line. After passing through the riddle box, the fruit continues on through a series of push pull extractors that remove any remaining waste material. The fruit then passes through the riffle washer, which consists of a stream of water over a series of angled baffles. The turbulence created causes the fruit to move over the baffles while leaving any heavier objects behind, it also cleans the fruit effectively. After being cleaned twice, the fruit is spun dried to remove any excess water. During the riffle washing stage however, the fruit absorbs some moisture [1].

### Final stage of processing and assessment

The Elbiscan laser scanner measures the reflectivity of the fruit as it passes through the scanner; objects that are outside the parameters are blown from the line, by a blast of compressed air while the fruit tumble over a drop on the conveyor belt. Trained staff inspects the fruit on picking tables before dressing

oil is applied to the fruit. The fruit is then packed to consumers' requirements in boxes with plastic liners. A liquid fumigant is added before the carton is sealed, this is followed by a final metal detection after which fruit is ready to be shipped to customers [1].

## **2.4 Grading**

Dried sultanas are graded on colour, from one to seven, based on the overall colour of the sample and the amount of variation in fruit colour. Primarily, the berries are separated according to Light and Brown type fruit and then further graded according to uniformity. Light type Seven crown grade, for example, are a uniform bright golden colour. Six crown grade are light amber with a maximum of 5% brown berries (no dark berries).

Five Crown grade have a maximum of 10% dark berries. Four crown fruit are amber coloured with a maximum of 15% dark berries. Three crown grade are berries of any shade of amber with a maximum of 20% dark berries. Two crown fruit, more than half are light coloured berries [11]

## **2.5 Factors affecting quality**

Several factors influence quality of processed sultanas, damage to the fruit caused by mishandling during harvest or damage due to the mechanical processing of the dried berries. There are several effects that reduce fruit quality, such as breaks in the skin that makes fruit sticky, causing the fruit to clump together. This is unappealing to the consumer as often there is a requirement for the fruit to easily break up and flow, particularly if it is a part of an automated food production system. Skin breaks in berries causes an increase in the rate of oxidative browning of the fruit making them darker in colour and therefore less desirable [13].

### **2.5.1 Enzymatic Browning**

There are numerous mechanisms of enzymatic browning which affect the quality of sultanas; one is the activity of *o*-diphenol oxidase, otherwise known as polyphenol oxidase (PPO) [10]. This mechanism is hindered by the use of cold dip treatments, as the rapid increase in sugar concentration retards the activity of this enzyme: with a more normal rate of drying, the sugar concentration changes more slowly allowing the browning to take place [14].

### **2.5.2 Sunlight Induced Browning**

Exposure to direct sunlight causes “sun burn” a natural breaking down of the chemicals within sultanas and results in an uneven darkening of the fruit. Such darkening can be avoided by the way the vines are pruned prior to fruit growth, and by using covered drying racks during drying [15].

### **2.5.3 Cultivar**

While the majority of fruit used in the production of dried sultanas are Thompson Seedless grapes, some variation in cultivar can affect quality. Merbein Seedless and Sultana H5 clones have been shown to give desirable features during drying with decreased drying times, which lessens the chances of enzymatic browning occurring, and better fruit colour compared to wild type cultivars. The Merbein Seedless variety (developed by CSIRO Merbein) while high yielding, produces fruit that are easily damaged when handled fresh. This makes them more suited to Trellis Drying, where the fruit are only handled dry [1].

### **2.5.4 Moisture Content**

During the drying process, if there is a change in humidity or rainfall that affects the crop, then the fruit can take up moisture. This increase in moisture content allows an increase in the activity of enzymatic browning, as enzyme activity is retarded by a high concentration of sugar within the fruit. This increase in moisture content can also increase the occurrence of mould and fungal growth [1].

### **2.5.5 Storage Time**

While the high concentration of sugars inhibits enzymatic browning, the reaction is only slowed, not halted. Over time, if the fruit is left in storage, browning will affect the fruit quality [16]. This effect can be minimised by storing the fruit at low temperatures, but using chemical means to further

inhibit the browning will cause the product to become less desirable to the consumer [17].

## **2.6 Analytical Procedures**

Several wet laboratory procedures have previously been used to determine chemical composition and nutritional value of foodstuffs, as shown in section 2.9. By using these tests in conjunction with NIR, it can be deduced whether there is correlation between features of the NIR spectra and results obtained with wet laboratory techniques:

### **2.6.1 Kjeldahl's Procedure, Crude Protein**

The organic material is digested in sulphuric acid in the presence of a catalyst that converts the nitrogen present into ammonium sulphate. The solution is then made basic and the resultant ammonia can then be steam distilled and titrated with hydrochloric acid to determine the protein content [18]. Refer to section 2.9.5 for examples of different applications of Kjeldahl protein combined with NIRS analysis.

### **2.6.2 Dew Point**

Heating a sample and allowing the resultant condensation to form on a cool mirror near the sample within the instrument, then measure the light reflected off that mirror allows a percentage of free (unbound) water to be determined within the sample. The amount of free water affects the shelf life of the dried fruit and whether further browning is likely. It also detrimentally influences the mechanical properties of the fruit and the likelihood of damage during

processing [16]. Examples of dew point analysis combined with and in conjunction with NIRS are given in section 2.9.2.

### **2.6.3 Chromameter**

Used to determine objectively, on the tri-stimulus scale of  $L^*$ ,  $a^*$  and  $b^*$  values, the colour of the object of interest. The Chromameter gives three values indicative of the colour of the object.  $L$  is the lightness scale, a value of 0 is black and a value of 100 indicates white. The Hue coordinates,  $a^*$  and  $b^*$ , are scaled from  $-60$  to  $60$  units. The  $a^*$  scale indicates from vivid green to vivid red and the  $b^*$  scale shows from vivid blue to vivid yellow respectively. Within the instrument, the data may be presented as either individual results or, as a series of measurements, then the mean and standard deviation of each colour parameter can be generated [16]. Some examples of CIE colour analysis combined with NIRS analysis are given in section 6.9.1.

### **2.6.4 Airstream sorting technique and grading**

Developed to assess the quality of naturally dried sultanas (“raisins” in the US) based on the proportion of full-bodied fruit within a sample. The airstream sorter operates by using a vertical flow of air of constant pressure and temperature, into which samples can be placed. The airflow pressure may be varied to allow each berry of the sample to drop through the column or be lifted and collected, depending on the airflow pressure and the weight of the fruit. The airflow pressure is expressed in inches of water [19, 20].

The fruit are graded depending on the percent of fruit that are either “blown” out of the column, at  $0.48$ ” of water, fruit that are blown are considered

substandard, and the percent that drop out at a pressure of 0.67” of water are considered B or better in grade. Proportion of C grade fruit are determined by the sum percent of substandard fruit and B or better fruit subtracted from 100 [19, 20].

### **2.6.5 Airstream sorting and maturity**

A series of projects have examined the link between maturity by Airstream sorting and other indicative parameters, such as colour, °brix at harvest titratable acidity and percent dry-mass [21].

Work by Kasimatis *et al* directly established a strong relationship between °brix and berry weight at harvest and percent B or better grades obtained from airstream sorting [20, 22], they compared clones of Thompson seedless fruit for both fresh and dried fruit performance with factors such as °brix, yield and airstream sorting grades.

Christansen *et al* presented work in two papers [23, 24] that undertook a three-year study into the influence of harvest date on Thompson seedless grapes, monitoring yield and quality of fresh and the resultant dried fruit. This study clearly demonstrated linear and curvilinear trends between the titratable acidity of dried fruit and harvest date and airstream sorting grades of the fruit.

## **2.7 Near infrared spectrophotometry**

### **2.7.1 Molecular spectroscopy theory**

Energy interacts with matter at the molecular level by causing vibrations of discrete frequency, the properties of which are related to the mass of the two atoms, in this case, a diatomic molecule, and the resistance of vibration due to the strength of the bond between them. This results in a distance versus energy diagram which shows the function of the force applied to the harmonic oscillator as a parabola, where an increase in distance requires an increase in energy applied to stretch the bond and likewise to compress the two atoms closer together [25]. This is valid for some diatomic molecules and can also give a good approximation for polyatomic molecules, however it makes several assumptions. In reality, the harmonic oscillator model is limited in its ability to predict the behaviour of vibrating molecules. Firstly, it assumes that the bounds of stretching a bond is unlimited, when in reality once a particular level of energy is reached and a distance is reached between the two atoms that there is no longer any sharing of electron density and the bond is considered broken. Likewise, the electron density holding the two atoms together also acts as a barrier that makes compression more difficult than assumed in the Harmonic Oscillator model [25, 26, 27].

### **2.7.2 Measurement and vibrational modes**

Vibrations can be measured by finding the telltale drop in energy intensity in the light that has passed through the sample, as the molecule has absorbed light at that frequency during vibrational excitation. A fundamental absorption

is when the energy difference between the relaxed state and the excited state is equal to one unit. Where a transition occurs from the ground to the excited state and the energy difference is greater than one, this is referred to as an overtone [25,26,27]. In a polyatomic system, there is more than one type of stretching deformation that the molecule can undergo, hybrids can also occur. For example, the molecule can undergo a stretching vibration while also bending. These vibrational modes are referred to as combinations [25,26,27].

### **2.7.3 Near Infrared Theory**

Infrared spectra arise from the fundamental vibrations of carbon, nitrogen and oxygen to hydrogen bonds and their respective overtones. As these bonds have a relatively low absorptivity within the near infrared region compared to mid Infrared, this fact allows the use of high concentrations of analyte using cells with longer path lengths, making the NIR spectroscopic technique ideal for the analysis of whole and unprocessed samples [28]. The implications of low absorptivity within the near infrared region can be most easily explained using the Beer-Lambert equation:

$$A = a b c$$

Because “a”, absorptivity is low within the near infrared region and “c”, the speed of light is constant then “b”, path length is relatively large compared to mid infrared spectroscopy. This gives NIRS the ability to examine samples without prior preparation or sample cells to contain the samples, allowing NIRS the ability to scan samples while they are passing by on the conveyor belt [28].

#### **2.7.4 Applications of near-infrared spectroscopy**

The development of NIR as an analytical tool has in comparatively recent times gained momentum, with the development of small powerful computers capable of handling the large amount of data received while processing NIR spectra. The increased use of NIR is also linked with the availability of techniques such as Chemometric analysis, allowing the relatively ambiguous NIR spectra to be more readily deduced. Near Infrared has been applied to the food industry for some years by fixed filter instruments capable of giving on-line readings of single factors such as protein, fat or moisture content, calibrated after more exhaustive testing by wet laboratory techniques that require invasive preparation and sampling techniques.

##### Food applications

Work by Bewig and co workers developed the use of NIR in conjunction with Mahalanobis distance principles, a form of statistical data handling, which showed that differentiation of vegetable oil types was possible and could be used as an alternative to Gas Chromatography. This was achieved by comparing the four different wavelengths of spectra of a number of samples of four different oil types. Firstly, a discriminate analysis model had to be developed using known samples. Nine samples of cottonseed oil, eight samples of soybean, eight peanut and five canola oil samples were recorded and compared to a series of samples, to validate the discrimination model. During the validation only two cases were misclassified, this was due to the assumption made by Mahalanobis Distances that the area occupied by each sample type is equivalent; the soybean group was shown to be more

dispersed than the others [29]. Edye and colleagues have found that the sugar concentration of both sugar beet and cane can be measured by the second derivative of the intensity of the spectra at a particular wavelength, compared to the concentration of sucrose reported by polarimetric measurements. This method has also been used to quantify the presence of Dextran, a by-product of bacterial infection in canes, which in high concentrations may cause problems in refining. This method is not as precise as more conventional methods, it can be used to quickly distinguish between high and low concentrations of dextran, thus reducing the number of samples that require more time consuming analysis [30].

#### Diffuse powdered samples

Komatsu and co-workers developed a MCFT-NIRS, a Multi-Channel Fourier Transform to measure diffuse powdered substances in reflectance mode. It was decided to use a Fourier Transform instrument, as this would allow the use of a small, relatively robust instrument, capable of quick analysis. Using a light source that would shine through a window built into the box containing the instrument, then onto the substance of interest which then receive reflected light from the sample, would make the instrument self-contained. The advantages of making the instrument self-contained are that the instrument is readily kept at ideal conditions. Hence, it is far more robust and flexible in its deployment and redeployment. This is due to the relative ease of relocating a self-contained instrument, that is therefore adaptable to the needs of the user [31].

### Whole Fruit samples

Schmilovitch and collaborators developed a way to determine both total soluble solids and moisture content by non-invasive NIR, such factors are of interest as they are indicative of future ripening. Whole dates were scanned by NIR then the spectra obtained were compared to results gained from an optical refractometer. Spectra were also compared with moisture content results obtained by measuring the mass of fruit before and after drying in a vacuum oven. Next, the data obtained was processed by chemometrics. By comparing the results obtained by NIR, the laboratory results and the parameters that show if fruit are likely to ripen, it was possible to discriminate “good dates” from “bad dates”. Fresh dates likely to ripen, contain less than 63% moisture and/or more than 32% soluble solids, those outside these parameters were defined as “bad dates”, as these fruit would not ripen [32].

## 2.8 Chemometrics

### 2.8.1 Multiple Linear Regression

Multiple linear regression splits the variables of experimental data into response variables and predictor variables, as can be seen in equation (i), below.

$$Y = b_0 + x_1b_1 + x_2b_2 + \dots \dots \dots (i)$$

The response variables “Y” being defined in this project as the data obtained by NIR spectrophotometry, while the predictor variables “x” are the results obtained by the wet laboratory techniques. There must be more predictor variables than response variables in the data set if MLR is to be used.

The performance of the model can be determined by dividing the data into two randomly chosen groups, then using one group to build the model and the second group to validate the model. Another option to validate the model is to randomly “leave one out”, then use the rest to build up the model, and compare with the results of the one left out. The model is valid in both cases if the validation data set, or the result left out, gives high correlation with the model data set [33].

### 2.8.2 Industrial and agricultural applications of MLR calibrations

Uddin and Okazaki evaluated NIR as a non-destructive technique to assess the authenticity of fresh fish in order to determine if they had been adulterated with frozen-thawed fish, which has a lower market value. Dry extract samples were assessed by NIR and samples originating from frozen-thawed fish were

found to have a lower absorbance. PCA and MLR were used to develop quantitative calibrations, it was, however, discriminate analysis techniques that proved capable of deducing the level of fresh and frozen-thawed fish [34]. Chen *et al* evaluated the effect of Multiplicative scatter correction on MLR calibrations of fat content in milk by comparing the performance of MLR calibrations using this pre-treatment, with untreated and second derivative spectra [35]. Maraboli *et al* investigated using NIRS to detect vegetable protein in milk powder. A series of samples of both genuine and adulterated milk powder were prepared varying from 0-5% added vegetable proteins. MLR using 5 wavelengths showed most promise in predicting vegetable protein content, achieving a correlation coefficient of determination of 0.993 [36].

Roggo and co-workers assessed NIRS as a replacement to test for sucrose in sugar beet instead of using a lead acetate procedure. A series of regression techniques were compared with the ability to update the calibration to taking into account variation between seasons, and the ability to transfer calibrations. It was found that a simple wavelength-by-wavelength calibration gave acceptable results while meeting the requirements of flexibility and future expansion [37]. Saranwong *et al* tested NIRS to evaluate whole mangoes for degrees Brix and dry matter using both MLR and PLS to develop the calibrations. Two separate regions (700-1100nm and 1100-2500nm) of the NIR spectra were also assessed for their associations with the constituents of interest. Both PLS and MLR showed promise in predicting degrees Brix and dry matter [38]. Ventura and co workers developed a calibration for assessing soluble solids in apples using MLR with a dual beam fibre optic

portable NIRS. This calibration demonstrated an ability to predict soluble solids with enough reliability to be used in the orchard [39]. Peiris and co workers prepared a series of calibrations for determining soluble solids content in individually processed tomatoes, using NIRS in transmittance mode [40].

### 2.8.3 Partial Least Squares

PLS is one of the most widely used calibration techniques in chemometrics. PLS uses the inverse calibration approach, which calibrates for desired components while accounting for other variations. PLS avoids the inversion problem by replacing original variables with linear combinations of the variables being examined [41,42]. Compared to MLR, PLS differs in several ways. The PLS statistics uses the whole spectrum which makes it more sensitive to outliers, at the same time PLS has the ability to detect features of an unusual shape thus indicating the validity of the model derived from the spectrum. Using the entire spectrum when developing a model, gave results of a lower standard deviation, in a similar manner to signal averaging.

Common to all inverse methods, including PLS, is how the relationship between concentrations and spectrophotometric measurements are modelled, as can be seen in the equation (ii) below:

$$c = R b \dots \dots \dots (ii)$$

where the vector “*c*” contains the known concentrations, “*R*” is a matrix of the measurements and “*b*” contains model coefficients. In PLS the “*R*” matrix is replaced with a new matrix, “*U*”, with columns that are linear combinations of the original columns of “*R*”. This new matrix uses both the variance and

covariance of the original matrix, while it has the same number of rows it has fewer columns, thus avoiding inversion problems. The relationship between the new matrix “*U*” and the matrix it is derived from “*R*” can be defined as:

$$U = R V S^{-1} \dots\dots(iii)$$

where “*V*” is the matrix containing loadings and “*S*” is a diagonal matrix of singular values. Principle components are then chosen. A minimum number is required to cover most of the variables of the model, however including too many may introduce instability when calculating “*S*<sup>-1</sup>” if PC’s with small diagonal elements are included. Now that the “*U*” matrix has been simplified to its truncated form “*U*<sup>T</sup>”, it is possible to solve for the regression vector “*b*<sub>reg</sub>”:

$$b_{est} = U^T c \dots\dots(iv)$$

where “*b*<sub>est</sub>” is the estimate of “*b*<sub>reg</sub>”. The resultant regression vector can now be used to predict the concentration of the unknown. Given the unknown has been measured by spectrophotometric analysis and using “*V*” and “*S*” from the calibration, the score vector can be obtained by the following equation.

$$U_{unknown} = c_{unknown} V S^{-1} \dots\dots(v)$$

Once “*U*<sub>unknown</sub>” is determined then with the previously determined “*b*<sub>est</sub>” it is used to determine the concentration of the substituent in the unknown sample using:

$$c = U_{unknown} b_{est} \dots\dots(vi).$$

[41,42].

#### **2.8.4 Industrial and agricultural applications of PLS calibrations**

A study was undertaken by Huang and co workers to examine the ability of NIR reflectance spectroscopy to assess foliar chemicals in eucalypt canopies using aircraft mounted spectrophotometers. Using continuum removal analysis techniques, on calibrations developed using PLS, of spectra of dried ground leaf samples, enabled applications of the technique to successfully analyse canopy foliage chemical properties [43]. Luypaert *et al* used NIRS coupled with PLS to determine the quality of green tea. A series of calibration models of both ground and whole leaves were developed to predict the content of caffeine, epigallocatechin gallate, epicatechin and total antioxidant capacity using the trolox equivalent antioxidant capacity method [44]. Delwiche and co workers used NIRS to assess whether related quality properties that occurred during wheat development were important. Twenty commercial cultivars were grown in ten geographical locations under the prevailing conditions to introduce some variation to the conditions in which the samples of the study were developed. Such a calibration would allow the identification of the variations in product development, at any point along the processing line [45]. Yang and Irudayargj examined a variety of spectroscopic techniques for the rapid analysis of vitamin C in foods and pharmaceutical products. Prediction reliability and analysis times for each method were compared [46]. Fardim *et al* used NIRS to assess a series of chemical and physico-chemical properties of unbleached eucalypt Kraft pulps, in diffuse reflectance mode, then used PLS to develop the model [47]. Cho and co-workers applied NIRS to the qualitative and quantitative analysis of velvet deer antlers. A quantitative model for ash content was developed using PLS

to replace the ignition technique, which is time consuming and cannot be undertaken *in-situ* [48].

## 2.9 Applications of NIRS

### 2.9.1 Colour and NIRS

#### Traditional , UV-Visible and NIR colour assessment

A number of projects have used NIRS to obtain spectra from the visible region of the spectrum to assess the subject of interest for colour. Such calibrations use colour quantification meters that employ systems such as the Hunterlab or CIE systems to quantify colour. The resulting data were used to develop calibrations in the same way as any other quantitative values and a linear regression equation is developed. One example of such a project is the work by Tsai *et al* where the soybean soak water was measured using a difference meter, recording the  $b^*$  value, amongst other parameters. This was done to develop a discriminant analysis calibration used to determine the season of the year the soybeans were grown [49].

#### Agricultural and Food Applications

Chen and Chen used near infrared transmission spectroscopy to develop a series of regression equations for edible vegetable oils. Both  $a^*$  (yellow), and  $b^*$  (red) were evaluated using a Nippon Denshoko ND-1001 DP colour meter. Stepwise regression analysis was used to develop the calibration and achieved correlation coefficients of 0.89 and 0.99, respectively [50]. Hammersly and Townsend developed a series of NIR reflectance calibrations to determine NIRS' ability to reliably predict colour measurements with the further application of using NIR to automatically assess scoured wool for

colour. Colour was assessed using a Hunterlab Colourquest and a Datacolour C53 using the CIE (1986) system [51]. Flinn and co workers developed calibrations for a wide variety of constituents for the quantitative assessment of Pulses, for both whole and ground samples. In general, ground samples were more accurately assessed for physical constituents such as colour, both field peas and chickpeas gave good calibrations for  $L^*$ ,  $a^*$  and  $b^*$  [52].

Shimatsu *et al* used NIT to evaluate physicochemical properties of both brown and milled rice, including Whiteness ( $L^*$ ) where the latter data was collected with a Minolta CR-200. The calibration was then developed using PLS and validated using full cross validation and test set validation [53]. Hong and Tsou used reflectance mode NIR to assess a number of quality parameters, including colour. Colour of the fruit was expressed as  $-a^*/b^*$  and MLR was used to develop the calibration [54]. Mc Caig used three different NIR spectrometers of the same manufacturer and model to measure colour values of a variety of agricultural products. CIE values were then obtained from the spectra using a standard formula developed by CIE and the values obtained were compared to colourimeter readings [55]. Leroy *et al* used PLS to develop a series of calibrations by means of a remote probe, in both transmission and reflectance modes, to rapidly assess beef samples that had undergone 2-8 days aging. The calibrations were developed for shear force of meat, drip loss and CIE colour [56].

Black and Panozzo used NIRS to reduce double handling of samples during analysis. An analytical technique capable of the simultaneous rapid

assessment of colour as well as other techniques was required: NIR proved ideal. Colour models based on CIE were developed with a number of other analytical techniques on several types of grains [57]. Liu *et al* studied the ability of predicting quality parameters of chicken breasts at different post-mortem times, using NIRS from 400-1850 nm. CIE colour, as well as other quality parameters were used to develop calibrations and showed increased ability to predict quality rather than individual sensory attributes [58]. Due to the ease that coffee can be unscrupulously contaminated, an accurate rapid analysis, that is of vital importance in quality assurance, was developed. Bogdanescu and co-workers used UV/Vis and NIRS spectroscopy for the rapid assessment of ground coffee. CIE and NIRS measurements of coffee containing varying levels of common contaminants were used to develop a discriminant analysis technique for rapid quality assessment [59].

## 2.9.2 Dew Point and NIRS

Interestingly few studies have been undertaken that use dew point values of subjects of interest, in conjunction with Near Infrared analysis, despite the widespread use of this technique throughout food and agricultural science. Most calibrations are used to assess parameters related to moisture use a percent moisture measurement, usually determined by oven drying to a constant mass, or other lengthy technique.

### Food applications

One exception is work by Huxsoll, which evaluated the ability of NIR to predict hydrated raisin dew point values. The calibrations developed were tested using cross validation, and exhibited potential in reliably assessing water activity of raisins [60]. Huxoll developed a series of calibrations of water activity for whole and ground seedless raisins and found that full spectrum calibrations could be used to reliably predict water activity of raisins. It was found, however, that vacuum oven moisture content could not be predicted accurately with the calibration developed [61]. Norris *et al* used NIR diffuse reflectance spectra of wheat starch and microcrystalline cellulose at  $a_w$  levels of 0.43 and 0.53 to determine whether NIR adsorption was more influenced by water activity or moisture content. It was found that moisture content could be distinguished and that adsorption spectra were not sensitive to different water activity levels at the same moisture content [62]. Hong and co workers used NIR spectra from 1100-2500 nm to measure unfreezable water bound to egg white lysosyme and soluble starch. Spectral features were found to be

attributable to freezable and unfreezable water, and have shown potential in determining unfreezable water content [63].

### **2.9.3 Lipids and NIRS**

As is the case for water activity there has only been limited attempts to use total lipids in NIRS calibrations. In many of these cases, the importance of this lipid constituent is secondary to individual fatty acids quantified by chromatographic techniques such as GC. Gonzales-Martin *et al* used NIRS in conjunction with solvent and microwave heating extraction, together with gas chromatography, to develop a series of calibrations. The calibrations developed could be used to determine the levels of twelve fatty acids, total poly- and mono-unsaturated as well as saturated fats by solvent extraction. Microwave heating was used on the same samples, which allowed the determination of six fatty acids, total poly- and mono-unsaturated and total saturated fats. Calibrations of all fatty acid parameters were used to develop calibrations using MPLS [64]. Gonzales-Martin and co-workers, combined their previous work in determining fatty acids in Iberian Breed Swine by using a remote fibre-optic probe, for the assessment of whole subcutaneous fat and solvent extracted fat. Calibrations were developed from this work using MPLS, which showed some utility in assessing fatty acids [65].

### **2.9.4 Titratable acidity and NIRS**

McGlone and co workers used NIR spectroscopy over 500-1100 nm wavelength range, to assess for pre- and post- harvest quality indices for Royal Gala apples, including titratable acidity. Chemometrics was used to

develop predictive models of these quality parameters, with some success [66]. Piers *et al* examined the utility of FT-NIRS reflectance spectroscopy, compared to more conventional NIRS techniques for the assessment of a number of quality parameters, including titratable acidity. They found that FT-NIR achieved a higher signal to noise ratio but had a lower penetration depth compared with dispersive NIRS. Both techniques, however, were able to measure the quality parameters assessed [67]. McGlone and co-workers examined different measurement modes of NIRS (reflectance, transmittance, Interactive) and spectral windows for their ability to predict harvest soluble solids and TA for Satsuma mandarins [68]. Navratil *et al* developed a technique for online assessment of yogurt and filmjolk fermentations using NIRS and electronic noise signals to assess amongst other parameters, pH and titratable acidity. PLS was used to develop the calibrations and the resultant work showed the potential of NIR, in conjunction with electronic noise data, as a means of rapid online assessment [69].

### **2.9.5 Protein and NIRS**

Due to the widespread use of Kjeldahl's procedure and the nature of this analysis, there has been a great deal of interest in developing rapid crude protein assessment in a wide variety of applications. Thus crude protein is one of the most common chemical parameters used to develop a NIRS calibration.

## Pasture and feed assessment

Andres and co workers assessed herbage from permanent meadows for nutritional value, namely dry matter, crude protein and ruminal degradability. Samples were fermented in cattle fitted with rumen cannulae and samples of both fermented and undigested samples were evaluated for potential and effective ruminal degradability. It was found that dry matter could be assessed with greater accuracy than crude protein due to contamination caused by microbial activity in the samples [70]. Confalonieri *et al* developed a series of dry matter constituents including crude protein, of natural alpine swards. These calibrations were developed using step-up, stepwise and MPLS techniques then compared and contrasted in performance. MPLS performed more precisely in dealing with chemical constituents, while stepwise was more effective for non-digestible fibre [71]. Bras and co workers investigated combining NIR and MIR spectra for use in analysing soybean flour quality parameters including crude protein. Calibrations were developed separately using PLS, the resultant calibrations were then combined using multi-block PLS and serial PLS, which showed that despite MIR's poor performance in isolation, it contained information that is absent in the NIR spectra [72].

Perez-Martin *et al* developed the prediction of chemical and ingredient composition of compound feedstuffs for different animals using NIRS. This experiment using both ground and unground samples to test the validity of using unground sample presentation for NIR [73]. Delwiche and Reeves studied the importance of pre-treatment in the development of PLS calibrations using a number of agricultural analytes such as crude protein on

wheat and forage samples [74]. The assessment of diet quality of free grazing animals is complicated by selective grazing, Borval and co-workers showed that faecal matter assessment with NIRS offered reliability in assessing dietary intake on various nutritive parameters, including crude protein, they then used MPLS to develop these calibrations [75].

Ciudad *et al* undertook a study to determine the ability of NIRS to assess the nutritive value of *Cystis multiflorus* as feed forage [76]. Fontaine *et al* assessed the repeatability of NIR spectrometers to measure amino acids analysis of feed raw materials [77]. White and Rouvinen-Watt evaluated the utility of NIRS to assess wet Mink feeds for various quality parameters, by developing calibrations for these parameters using MPLS [78]. Alomar *et al* examined the changes to spectral features and chemical composition of forage samples due to a number of sample drying techniques. The analysed results were further refined using PCA and it was found that freeze drying of samples had little effect, oven drying had notably decreased soluble crude protein and digestible organic matter in feeds, but insoluble nitrogen and neutral detergent fibre increased [79]. Paul and co workers assessed NIRS for its ability to provide continuous quality control in forage harvesters, with an emphasis on the effect of temperature variation on NIR adsorption on a number of calibrations including crude protein, to establish the importance of temperatures influence on prediction. Using a broad range of temperatures during the development of a calibration, the problem can be largely avoided [80].

Alomer *et al* examined the effects of drying methods on chemical composition and NIR spectra of pasture silage, by studying a number of quality parameters including Kjeldahl protein. The results were then examined using discrete analysis. The conclusions made indicated sample preparation should be constant throughout the development of NIR calibrations [81]. Uray *et al* compared NIRS to classical methods for determining fodder quality including crude protein [82]. Alomar *et al* examined the effect of sample preparation and drying techniques on spectral features and chemical composition of fresh silage [83].

#### Crop assessment

Gatius *et al* analysed dried and milled samples of wheat for protein content to discriminate samples according to maturity. PCA was used for discriminant analysis and showed ability to distinguish between growth stages of wheat samples. The crude protein model was developed two ways, by using a “global model” consisting of all sample types in the study, and separate “local models” made up of separate types of samples. The “global model” proved more reliable [84]. Fassio and Cozzolino investigated NIRS as a technique for assessing a series of quality parameters in whole sunflower seeds, including crude protein. MPLS was used to develop the resultant calibrations, which showed sufficient reliability for pre-screening of samples [85]. Mika and collaborators developed and assessed whole oil rapeseed calibrations for their ability to assess a series of constituents, including crude protein [86]. Kays developed a series of calibrations to assess mixed grain cereals for a number of quality parameters including protein. The obtained calibration was

developed using MPLS and successfully achieved the precision required by US nutritional labelling [87]. Bruno-Soares and co workers used NIRS to assess a variety of green crop grain quality parameters including protein. NIRS calibrations showed acceptable reliability in predicting composition across different species of green crop cereals [88].

### Food analysis

Curda and Kukackova used FT-NIR with a fibre optic probe to assess its ability to measure a variety of parameters of processed cheese, then developed a series of calibrations using PLS [89]. Purnmoadi and co-workers investigated the effect of feed source on milk quality parameters, including protein. The study demonstrated that fat content was not significantly affected but protein content was affected by feeding regimes [90]. Laporte and Paguin used NIT to determine a number of quality parameters of milk, including crude and true protein. Calibrations were developed using PLS for both, homogenised, local and, homogenised and non-homogenised, global calibrations. In almost all cases, the global calibrations proved more reliable than the local calibrations, and despite the limited size of the calibration, the NIR proved reliable in predicting milk quality parameters [91]. Buning-Pfaue and collaborators developed three methods of quantitative assessment of high moisture content foods including fat, crude protein and carbohydrates [92].

## Agriculture

Woo *et al* examined the potential for NIRS to classify the cultivation area of Ginseng from Korea and China, using SIMCA to undertake discriminant analysis on a variety of components including crude protein. The key factors for determining region of origin were found to be starch and the inorganic elements present in the samples [93]. Berardo *et al* examined the viability of assessing key Pigeon Pea quality parameters using NIRS and achieved good correlation coefficients [94].

## Laboratory quality assurance

Ruisanchez and co workers undertook a proficiency study of seven participating laboratories, in using their preferred multivariate analysis technique and software, on a set of data using commonly analysed components such as crude protein. The same data set was sent to the contributing groups and the resultant calibrations were compared [95].

Bakalli *et al* compared results obtained using manufacturer supplied calibration curves, to results obtained from the same samples assessed using AOAC procedures, including crude protein. The study concluded, it would be of value to run known standards and adjust the result for unknowns by the observed deviations from the standards [96].

## **2.10 Aims**

### **2.10.1 General aims of the project**

This project has the general aim to investigate the uses of NIR as a more precise method for monitoring the quality of sultanas. This analysis will be applied to the final stages at the processing line in place of the current Crown Grade Classification system.

### **2.10.2 Specific aims of the project**

- (i) To determine the quality parameters: colour, water activity, Kjeldahl protein and titratable acidity by modification/adaptation of literature procedures and to then correlate these data with the NIR spectra of processed fruit.
- (ii) Develop calibrations to assess the maturity of unprocessed fruit, by the application of calibrations developed from Titratable acidity of a diverse set of samples separated using Airstream Sorting, which will ensure a wide range of samples of varying maturity.
- (iii) To develop and validate calibrations for the above quality indicators by selecting suitable visible and NIR wavelengths, employing several statistical analysis methods, then undertake prediction tests with other samples.

## 2.11 Bibliography

1. *ADFA Dried Vine Fruit Manual*. Australian Dried Fruits Association inc. 1998
2. *ADFB Annual Report, 1999-2000*.
3. Macrae, I. "The Production of Natural Sultanas", *Dried fruit news*. 1989 February: 8-9
4. Kerridge, GH and Grncarevic, M. "Production of Dark Coloured Sultanas by Sun Drying in Australia", *Food Technology in Australia* 1965; 17: 328-331.
5. Grncarevic, M. Lewis, WJ. "Drying of grapes in Australia", *Food Technology in Australia* 1976 28: 66-67, 69-71,76.
6. Fuller, RJ. "The Origin and Development of the Australian Grape Drying Rack", *Dried Fruit News* 1989 December: 13-15.
7. Fuller, RJ, Schache, MJ, Kaye, DR. "Improved Technology for Solar Drying of Vine Fruit", *Department of Agricultural And Rural Affairs Research Report Series* 1989 June; No. 86: Agdex 243-55.
8. Gould, IV and Whiting, JR. "Mechanisation of Raisin Production with the Irymple Trellis System" *Transactions of the ASAE* 1987 January-February; 30 (1): 56-60.
9. Clingeffer, PR. "Sultana raisin production". In: Petrucci, VE. Clary, CD. Editors. *A Treatise on Raisin Production, Processing and Marketing*. Clovis, California. :Malcolm Media Press, 2002: 131-144.
10. Uhlig, BA. "The Production of Light Coloured Dried Sultanas", *Die Weinwissenschaft: viticultural and enological sciences* 1996; 51(1): 37-39.

11. Grncarevic, M, Lewis, W. "External Colour of Dried Sultanas", *Food Technology in Australia* 1973 November: 562-565.
12. Uhlig, BA, Clingeleffer, PR. "Ripening Characteristics of the Fruit from *Vitis Vinifera* L. Drying Cultivars Sultana and Merbien Seedless under Furrow Irrigation", *American Journal of Enology and Viticulture* 1998; (49) 4: 375-382.
13. Harris, JM Grncarevic, M. " A Method for the Estimation of Damage on Dried Sultanas", *Food Technology in Australia* 1968 October: 472.
14. Grncarevic, M, Hawker, JS. "Browning of Sultana Grape Berries During Drying", *Journal of the Science of Food and Agriculture* 1971 May; Vol. 22: 270-272.
15. Uhlig, BA. "Effects of solar Radiation on Grape (*Vitis Vinifera* L.) Composition and Dried Fruit Colour" *Journal of Horticultural Science and Biotechnology* 1998; 73(1):111-123.
16. Frank, D. *Investigation of the biochemical basis of browning during the storage of sultanas, a PhD Thesis*. Victoria University, Werribee Campus; 2001.
17. Mc Bride, RL, Mc Bean, DMcG, Kuskis, A. "The Shelf Life of Sultanas: a Sensory Assessment" *Lebensmittel-Wissenschaft Und-Technologie* 1984;17:134-136.
18. *Vogel's Textbook of Quantitative Chemical Analysis*, Fifth Edition. New York: Longman Scientific and Technical; 1989.
19. Anonymous. *Official Methods of Analysis*. 6<sup>th</sup> edition. AOAC. Fruit and fruit products. 1945; p.596

20. Kasimatis, AN, Vilas, EP Jr., Swanson, FH, Baranek, PB. "Relationship of Soluble Solids and Berry Weight to Airstream Grades", *American Journal of Enology and Viticulture* 1977; 28(1): 8-15.
21. Uhlig, BA, Clingeleffer, PR. "Influence of Grape (*Vitis Vinifera* L.) Berry Maturity on Dried Fruit Colour" *Journal of Horticultural Science and Biotechnology* 1998; 73(3): 329-339.
22. Christensen, LP, Bianchi, ML. "Comparison of Thompson Seedless Clones for Raisin Production", *American Journal of Enology and Viticulture* 1994; 45(2): 150–154.
23. Christensen, LP, Bianchi, ML, Lynn, CD, Kasimatis, AN, Miller, MW. "The Effects of Harvest Date on Thompson seedless Grapes and Raisins (I) Fruit Composition, Characteristics and Yield", *American Journal of Enology and Viticulture* 1995; 46(1):10-16.
24. Christensen, LP, Bianchi, ML, Miller, MW, Kasimatis, AN, Lynn, CD. "The Effects of Harvest Date on Thompson seedless Grapes and Raisins (II) Relationships of Fruit Quality Factors", *American Journal of Enology and Viticulture* 1995; 46(4):493-498.
25. Struve, WS. *Fundamentals of Molecular Spectroscopy*. Oxford Chemistry Primer. New York: Wiley Interscience; 1989.
26. Miller, CE. "Chemical principles of NIR technology." In: Williams, PC, Norris, K, editors. *Near Infrared Technology in the Agricultural and Food Industries*, 2<sup>nd</sup> Edition. St. Paul (Mi): American Association of Cereal Chemists; 2001. p. 19-37.

27. Hertzburg, G. *Molecular spectra and structure. II Infrared and Raman Spectra of Polyatomic Molecules*. D. Van Nostrand Co., Inc., Princetown, NJ. 1950.
28. Osborne, BG, Fearn, T, Hindle, PH. *NIR spectroscopy with applications in food and beverage analysis*. 2<sup>nd</sup> Edition. Singapore: Longman Scientific and Technical; 1993.
29. K. M. Bewig, A. D. Clarke, C. Roberts, N. Unklebay. "Discriminant analysis of vegetable oils using near infrared spectroscopy." In: Batten, GD, Flinn, PC, Welsh, LA, Blakeney, AB, editors. *Leaping ahead with Near Infrared Spectroscopy*. Melbourne: NIRSG, RACI; 1995; pp 324-326.
30. L. A. Edye, M. A. Clarke, C. McDonald-Lewis. "Application of NIR in Sugar Cane and Sugar Beet Processing and Sugar Manufacture." In: Batten, GD, Flinn, PC, Welsh, LA, Blakeney, AB, editors. *Leaping ahead with Near Infrared Spectroscopy*. Melbourne: NIRSG, RACI; 1995; p. 340-343.
31. T. Komatsu, H. Ueharu, S. Shinkai, K. Sakai. "Development of an On-Line Constituent Analysis/Judgement System for the Flour Milling Process." In: Batten, GD, Flinn, PC, Welsh, LA, Blakeney, AB, editors. *Leaping ahead with Near Infrared Spectroscopy*. Melbourne: NIRSG, RACI; 1995; p. 334-339.
32. Z. Schmilovitch, Z. Bernsien, M. Austerweil, A. Zaltzman, G. D. Dull. "Fresh Date Sorting by NIRS." In: Murray, I, Cowe, IA, editors. *Making light work: Advances in near infrared spectroscopy*. New York: Ian Michael Publications; 1992; p. 425-429.

33. Miller, JN, Miller, JC. *Statistics and Chemometrics for Analytical Chemistry*. 5th Edition. Sydney: Prentice Hall; 2000.
34. Uddin, M, Okazaki, E. "Classification of Fresh and Frozen Fish by Near Infrared Spectroscopy" *Journal of Food Science* 2004 October; 69(8):C665-C668.
35. Chen, JY Iyo, C, Terada, F, Kawano, S. "Effect of Multiplicative Scatter Correction on wavelength selection for Near Infrared Calibration to Determine Fat Content in Raw Milk", *Journal of Near Infrared Spectroscopy* 2002; 10(4): 301-307.
36. Maraboli, A, Cattaneo, TMP, Giangaicomo, R. "Detection of Vegetable Proteins from Soy, Pea and Wheat Isolates in Milk Powder by Near Infrared Spectroscopy", *Journal of Near Infrared Spectroscopy* 2002;10(1):63-69.
37. Roggo, Y, Duponchel, L, Noe, B, Huvenne, JP. "Sucrose Content Determination of Silver Beets by Near Infrared Reflectance Spectroscopy" *Journal of Near Infrared Spectroscopy* 2002; 10(2):137-150.
38. S. Saranwong, J. Sornsrivichai, Kawano. S. "Improvement of PLS Calibration for Brix Value and Dry Matter of Mango using Information from MLR Calibration", *Journal of Near Infrared Spectroscopy* 2001; 9(4): 287-295.
39. Ventura, M, de Jager, A, de Putter, H, Roelofs, FPMM. "Non-destructive determination of Soluble Solids in Apple Fruit By Near Infrared Spectroscopy (NIRS)", *Postharvest Biology and Technology* 1998 September; 14(1):21-27.

40. Peiris, KHS, Dull, GG, Leffler, RG, Kays, SJ. "Near Infrared (NIR) Spectrometric Technique for Non-destructive Determination of Soluble Solids Content in Processing Tomatoes", *Journal of the American Society for Horticultural Science* 1998; 23(6):1089-1093.
41. Adams, MJ. *Chemometrics in Analytical Spectroscopy*. Cambridge: The Royal Society of Chemistry; 1995.
42. Beebe, KR, Pell, RJ, Seasholtz, MB. *Chemometrics: A Practical Guide*. Brisbane: J Wiley and Sons; 1998.
43. Huang, Z, Turner, BJ, Dury, SJ, Wallis, IR, Foley, WJ. "Estimate Foliage Nitrogen Concentration from HYMAP Data by Continuum Removal Analysis", *Remote Sensing of Environment* 2004 October 30;93(1-2):18-29.
44. Luypaert, J, Zhang, MH, Massart, DL. "Feasability Study for the Use of Near Infrared Spectroscopy in the Qualitative and Quantitative Analysis of Green Tea, *Camellia Sinensis* (L.) *Analytica Chimia Acta* 2003; 478 (2):303-312.
45. Delwiche, SR, Graybosch, RA, Nelson, LA, Hruschka, WR. "Environmental Effects on Developing Wheat as Sensed by Near-Infrared Reflectance of Mature Grains", *Cereal Chemistry* 2002 Nov-Dec; 79(6): 885-891.
46. Yang, H, Irudayaraj, J. "Rapid Determination of Vitamin C by NIR, MIR and FT-Raman Techniques", *Journal of Pharmacy and Pharmacology* 2002 September; 54(9):1247-1255.

47. Fardim, P, Ferreira, MMC, Duran, N. "Multivariate Calibration for Quantitative Analysis of Eucalypt Kraft Pulp by NIR Spectrometry", *Journal of Wood Chemistry and Technology* 2002; 22(1):67-81.
48. Cho, CH, Woo, YA, Kim, HJ, Chung, YJ, Chang, SY, Chung H. "Rapid Qualitative and Quantitative Evaluation of Deer Antler (*Cervus elaphus*) Using Near Infrared Reflectance Spectroscopy", *Microchemical Journal* 2001 March;68(2-3):189-195.
49. Tsai, S-J, Hong, TL Tsou, SCS. "Discrimination in the Processing Quality of Soybeans Harvested at Different Seasons by NIRS." In: Batten, GD, Flinn, PC, Welsh, LA, Blakeney, AB, editors. *Leaping ahead with Near Infrared Spectroscopy*. Melbourne: NIRSG, RACI; 1995; p218-221.
50. Y. S. Chen, A. O. Chen. "Quality analysis and purity examination of edible vegetable oils by NIT spectroscopy". In: Batten, GD, Flinn, PC, Welsh, LA, Blakeney, AB, editors. *Leaping ahead with Near Infrared Spectroscopy*. Melbourne: NIRSG, RACI; 1995; p.316-323.
51. M. J. Hammersly, P. E. Townsend. "Exploiting the Shorter Wavelengths: Wool Measurements Including Colour". In: Batten, GD, Flinn, PC, Welsh, LA, Blakeney, AB, editors. *Leaping ahead with Near Infrared Spectroscopy*. Melbourne: NIRSG, RACI; 1995; p.465-469.
52. Flinn, PC, Black, RG, Iyer, L, Brouwer, JB, Meares, C. "Estimating the Food Processing Characteristics of Pulses by Near Infrared Spectroscopy, using Ground or Whole Samples", *Journal of Near Infrared Spectroscopy* 1998;6:213-220.

53. Shimatsu, N, Katsura, J, Yanagisawa, T, Tezuka, B, Maruyama, Y, Inoue, S, Eddison, CG, Cowe, IA, Withey, RP, Blakeney, AB, Kimura, T, Yoshizaka, S, Okadome, H, Toyoshima, H, Ohtsubo, K. "Evaluating Techniques for Rice Grain Quality Using Near Infrared Transmission Spectroscopy", *Journal of Near Infrared Spectroscopy* 1998;6:A111-A116.
54. Hong, TL, Tsou, SCS. "Determination of Tomato Quality by Near Infrared Spectroscopy", *Journal of Near Infrared Spectroscopy* 1998;6:A321-A324.
55. Mc Caig, TN. "Extending the Use of Visible/Near-Infrared Reflectance Spectrophotometers to Measure Colour of Food and Agricultural Products", *Food Research International* 2002;35(8):731-736.
56. Leroy, B, Lambotte, S, Dotreppe, O, Lecocq, H, Istasse, L. "Prediction of Technological and Organoleptic Properties of Beef Longissimus Thorasis from Near Infrared Reflectance and Transmission Spectra", *Meat Science* 2003;66(1):45-54.
57. Black, CK, Panozzo, JF. "Accurate Techniques for Measuring Colour Values of Grain and Grain Products using a Visible-NIR Instrument" *Cereal Chemistry* 2004;81(4):469-474.
58. Liu, Y, Lyon, BG, Windham, WR, Lyon, CE, Savage, EM. "prediction of Physical Colour and Sensory Characteristics of Broiler Breasts by Visible-NIR Reflectance Spectroscopy", *Poultry Science* 2004;83(8):1467-1474.

59. Bogdanescu, V, Giurginca, M, Iftimie, N, Mihalache, R, Meghea, A, *et al.* "The Determination of Impurities from Coffee by Spectral Methods", *Revista de Chimie* 2005;56(4):378-381.
60. Huxoll, CC. "Near Infrared Analysis Potential for Measuring the Water Activity of Raisins". *Conference Proceedings of the IFT Annual Meeting* 1995 p70.
61. Huxoll, CC. "Assessment of Near Infrared (NIR) Diffuse Reflectance analysis for measuring Moisture and Water Activity in Raisins", *Journal of Food Processing and Preservation* 2000;24(4): 315-333.
62. Delwiche, SR, Pitt, RE, Norris, KH. "Sensitivity of Near Infrared Adsorption to Moisture Content Versus Water Activity in Starch and Cellulose", *Cereal Chemistry* 1992;69(1); 107-109.
63. Hong, J, Yamaoka-Koseki, S, Tsujii, Y, Yasumoto, K. "Applicability of Near Infrared Spectroscopic Method to Unfreezable Water Measurements in Egg White Lysozyme and Soluble Starch", *Liebensmittel-Wissenschaft Und-Technologie* 1997;30(4):406-410.
64. Gonzales-Martin, I, Gonzales-Perez, C. Hernandez-Mendez, J, Alvarez-Garcia, N, Lazaro, SM. "Determination of Fatty Acids in the Subcutaneous Fat of Iberian Breed Swine by Near Infrared Spectroscopy. A Comparative Study of the Methods for Obtaining Total Lipids: Solvents and Melting With Microwaves", *Journal of Near Infrared Spectroscopy.* 2002;10(4):257-268.
65. Gonzales-Martin, I, Gonzales-Perez, C. Hernandez-Mendez, J, Alvarez-Garcia, N. "Determination of Fatty Acids in the Subcutaneous

- Fat of Iberian Breed Swine by Near Infrared Spectroscopy (NIRS) With a Fibre Optic Probe”, *Meat Science* 2003;65(2):713-719.
66. McGlone, VA, Jordan, RB, Martinsen, PJ. “Visible-NIR Estimation at Harvest of Pre- and Post-storage Quality Indices for Royal Gala Apple”, *Post-harvest Biology and Technology* 2002;25(2):135-144.
  67. Piers, A, Scheerlinck, N, Touchant, K, Nicolai, BM. “Comparison of Fourier Transform and Dispersive Near Infrared Reflectance Spectroscopy for Apple Quality Measurements”, *Biosystems Engineering* 2002;81(3):305-311.
  68. McGlone, VA, Fraser, DG, Jordan, JB, Kunemeyer, R. “Internal Quality Assessment of Mandarin Fruit by Visible-Near Infrared Spectroscopy”, *Journal of Near Infrared Spectroscopy* 2003;11(5):323-332.
  69. Navratil, M, Cimander, C, Mandenius, CF. “Online Multi-Sensor Monitoring of Yoghurt and Filmjolk Ferments on Production Scale”, *Journal of Agricultural and Food Chemistry* 2004;52(3):415-420.
  70. Andres, S, Calleja, A, Lopez, S, Mantecon, AR, Giraldez, FJ. “Nutritive Evaluation of Herbage from Permanent Meadows by Near Infrared Reflectance Spectroscopy (II): Prediction of Crude Protein and Dry Matter Degradability”, *Journal of the Science of Food and Agriculture* 2005;85(9):1572-1579.
  71. Confalonieri, M, Lombardi, G, Bassignana, M, Odoardi, M. “Analysis of Quality Constituents of Natural Alpine Swards with Near Infrared Reflectance Spectroscopy”, *Journal of Near Infrared Spectroscopy* 2004;12(6):411-417.

72. Bras, LP, Bernardino, SA, Lopes, JA, Menezes, JC. "Multiblock PLS as an Approach to Compare and Combine NIR and MIR Spectra in Calibrations of Soybean Flour", *Chemometrics and Intelligent Laboratory Systems* 2005;75(1):91-99.
73. Perez-Martin, DC, Garrido-Varo, A, Guerrero-Ginel, JE, Gomez-Cabrera, A. "Near Infrared Reflectance Spectroscopy (NIRS) for the Mandatory Labelling of Compound Feeding-stuffs: Chemical Composition and Open-Declaration", *Animal Feed Science and Technology* 2004;116(3-4):333-349.
74. Delwiche, SR, Reeves, JB. "The Effect of Spectral Pre-treatments on the Partial Least Squares Modelling of Agricultural Products", *Journal of Near Infrared Spectroscopy* 2004;12(3):177-182.
75. Borval, M, Coates, DB, Lecomte, P, Decruyenaere, V, Archimede, H. "Faecal Near Infrared Reflectance Spectroscopy (NIRS) to Assess Chemical Composition *In Vivo* Digestibility and Intake of Tropical Grass by Creole Cattle", *Animal Feed Science and Technology* 2004;114(1-4):19-29.
76. Ciudad, AG, Santos, BF, de Aldana, BRV, Zabalgoageazcoa, I, Gutierrez, MY, Criado, BG. "Use of Near Infrared Reflectance Spectroscopy to Assess Forage Quality of a Mediterranean Shrub", *Communications in Soil Science and Plant Analysis* 2004;35(5-6):665-678.
77. Fontaine, J, Horr, J, Schirmer, B. "Amino Acid Contents in Raw Materials can be Precisely in a Global Network of Near Infrared Spectrometers: Collaborative Trials Prove the Positive Effects of

- Instrument Standardisation and Repeatability Files”, *Journal of Agriculture and Food Chemistry* 2004;52(4):701-708.
78. White, M, Rouvinen-Watt, K. “Near Infrared Evaluation of Wet Mink Diets”, *Animal Feed Science and Technology*. 2004;111(1-4):239-246.
79. Alomar, D, Fuchslocher, R, de Pablo, M. “effect of preparation Method on Composition and NIR Spectra of Forage Samples”, *Animal Feed Science and Technology* 2003;107(1-4):191-200.
80. Paul, C, Dietrich, F, Rode, M. Land “Influence of Sample Temperature on the Assessment of Quality Characteristics in Undried Forages by Near Infrared spectroscopy (NIRS)”, *Bauforschung Volkenrode* 2002;52(4):229-237.
81. Alomar, D, Fuchslocher, R, Stockebrand, S. ‘Effects of Oven- or Freeze-Drying on Chemical Composition and NIR Spectra of Pasture Silage”, *Animal Feed Science and Technology* 1999;80(3-4):309-319.
82. Uray, G, Kaufmann, J, Steiner, B. “Quality Assurance of the Near Infrared Spectroscopy in Fodder Analysis: Determination of Dry Matter, Crude Protein, Crude Fibre and Crude Fat”, *Bodenkultur* 1999;50(1):25-38.
83. Alomar, D, Montero, R, Fuchslocher, R. “Effect of Freezing and Grinding Method on Near-Infrared Reflectance (NIR) Spectra Variation and Chemical Composition of Fresh Silage”, *Animal Feed Science and Technology* 1999;78(1-2):57-63.
84. Gatius, F, Lloveras, J, Ferran, J, Puy, J. “Prediction of Crude Protein and Classification of the Growth Stage of Wheat Plant Samples From NIR Spectra”, *Journal of Agricultural Science* 2004;142:517-524.

85. Fassio, A, Cozzolino, D. "Non-destructive Prediction of Chemical Composition in Sunflower Seeds by Near Infrared Spectroscopy", *Industrial Crops and Products* 2004;20(3):321-329.
86. Mika, V, Tillmann, P, Koprna, R, Nerusil, P, Kucera, V. "Fast Prediction of Quality Parameters in Whole Seeds of Oilseed Rape (*Brassica Napus L.*)", *Plant Soil and Environment* 2003;49(4):141-145.
87. Kays, SE, Barton, FE, Windham, WR. Predicting Protein Content by Near Infrared Reflective Spectroscopy in Diverse Cereal Food Products", *Journal of Near Infrared Spectroscopy* 2000;8(1):35-43.
88. Bruno-Soares, AM, Murry, I, Paterson, RM, Abreau, JMF. "Use of Near Infrared Reflectance Spectroscopy (NIRS) for the Prediction of the Chemical Composition and Nutritional Attributes of Green Crop Cereals", *Animal Feed Science and Technology* 1998;75(1):15-25.
89. Curda, L, Kukackova, O. "NIR Spectroscopy: A Useful Tool for Rapid Monitoring Processed Cheeses Manufacture", *Journal of Food Engineering*. 2004;61(4):557-560.
90. Purnmoadi, A, Batajoo, KK, Ueda, K, Terada, F. "Influence of Feed source on Determination of Fat and Protein in Milk by Near infrared Spectroscopy", *International Dairy Journal* 1999;9(7):447-452.
91. Laporte, MF, Paguin, P. "Near infrared Analysis of Fat, Protein and Casein in Cows' Milk" *Journal of Agricultural and Food Chemistry* 1999;47(7):2600-2605.
92. Buning-Pfaue, H, Hartmann, R, Harder, J, Kehraus, S, Urban C. "NIR-Spectrometric Analysis of Food. Methodical Development and

- Achievable Performance Values”, *Fresenius Journal of Analytical Chemistry* 1998;360(7-8):832-835.
93. Woo, YA, Cho, CH Kim, HJ, Yang, JS, Seong, KY. “Classification of Cultivation Area of Ginseng by Near Infrared Spectroscopy and ICP-AES”, *Microchemical Journal* 2002;73(3):299-306.
94. Berardo, N, Dzwola, BH, Hove, L, Odoardi, M. “Near Infrared Calibration of Chemical Constituents of *Cajanus cajan* (Pigeon Pea) Used as Forage”, *Animal Feed Science and Technology* 1997;69(1-3):201-206.
95. Ruisanchez, I, Rius, F, Maspoch, S, Coello, J, Azzouz, T, Tauler, R, Sarabia, L, Ortiz, MC, Fernandez, JA, Massart, D, Puigdomenech, A, Garcia, C. “preliminary Results of an Interlaboratory study of Chemometric Software and Methods on NIR Data Predicting the Content of Crude Protein and Water in Forages”, *Chemometrics and Intelligent Laboratory Systems* 2002;63(2):93-105.
96. Bakalli, RI, Pesti, GM, Etheridge, RD. “Comparison of a Commercial Near Infrared Spectroscopy and Standard Chemical Assay Procedures for Analysing Feed Ingredients: Influence of Grinding Methods”, *Journal of Applied Poultry Research* 2000;9(2):204-213.

## **3.0 Materials and Methods**

### **3.1 Introduction**

Quality assessment of fruit has typically been carried out only at the transition between grower and packer, as the grower is paid according to the quality of the dried fruit at receipt. Unfortunately, there are two major problems with this system, firstly the assessment itself lacks objectivity and is based purely on the fruit colour, and secondly a packer also has to sell a product to meet a buyer's requirements, whose expectations of a product are based on an assessment that was limited in scope, subjective and of a parameter that often degrades with time. A system is required where an assessment of the processed product is made as it leaves the production line. This should more readily ensure the requirements of a customer are more readily met by the packer's product. The assessments possible using NIRS are greatly increased beyond a colour assessment, as new parameters can be determined rapidly, assuming a correlation can be found between features on NIRS spectra and analyte levels in fruit.

## **3.2 Part I: NIRS assessment of processed sultanas**

### **3.2.1 Sampling Procedure**

Samples were collected at Sunbeam Foods, Mildura, Sarnia Packing, Irymple and Angus Park in South Australia (the latter two companies now owned by Sunbeam Foods). Samples were collected during processing just prior to boxing, in either hourly or half hourly intervals. A general aim of sample collection was that samples should be as diverse as possible. Only sultanas were to be collected, either “golden” or “natural” type sultanas. Samples were typically between 200 and 500 grams in size and collected in batches of forty to fifty samples, throughout the year.

### **3.2.2 Sample Preparation**

As the fruit has undergone processing, cleanup of samples is not required since all vine trash, empty husks and leaf matter has long gone from the sample. Samples once received from Mildura were stored in the laboratory under ambient conditions, it was felt that keeping them in cold storage would alter their moisture content and cause crystallisation of the sugars within the fruit. Due to the nature of the product, the fruit were analysed in batches. Samples were analysed for colour, water activity and titratable acidity within three days of NIRS analysis, and samples were prepared and pre-weighed for Kjeldahl’s procedure and total lipids analysis. For the destructive techniques that require homogenation, a representative sample of ~30 grams was placed into a mortar and ground with a pestle to a paste-like consistency. This was sufficient for the analytical methods used on one sample and would ensure that the sample was representative.

### **3.2.3 Specific preparation for Titratable acidity experiment**

Approximately 10 grams of ground sample was added to a tared 500ml single necked round-bottomed flask, this was then accurately weighed and recorded. 50ml of MiliQ water (Milipore MiliQ Plus) was then added to the flask and the fruit paste was left to hydrate for 30 minutes. A reflux condenser was fitted to the flask and then brought to reflux for 90 minutes using a heating mantle. After cooling, the resultant mixture was then transferred to a 200ml volumetric flask with rinsing, including pulp, and then made to 200ml with MiliQ water (Milipore MiliQ Plus). The contents were then gravity filtered through a Watmann grade #1 qualitative filter paper into a 200 ml flat bottle. The filtered solution was then stored at 4<sup>0</sup> C while it awaited titration [1].

### **3.2.4 Specific preparation for Kjeldahl's procedure**

Approximately 2.5 grams of ground sample was weighed into a pre tared 70ml grade #1 Watmann filter paper and accurately weighed. A duplicate sample was also prepared. Each sample was wrapped up with the filter paper and placed into a labelled plastic resealable bag and then stored at 4<sup>0</sup> C until required for analysis [2].

### **3.2.5 Specific preparation for total lipids determination**

Approximately 5 grams of ground sample was accurately weighed into a tared 50 ml falcon tube. A duplicate was also prepared. Each sample and duplicate were appropriately labelled and stored at 4<sup>0</sup> C until required [3].

### **3.3 Experimental Analysis**

#### **3.3.1 CIE tricolour stimulus**

Processed whole berries were tightly packed into a 200ml mortar until full. Ten measurements using a Milolta CR-310 Chromameter were then taken across the upper surface of the sample, the mortar was then emptied and repacked and another ten measurements were taken of the sample. This process was continued until a total of 50 measurements were taken. The average  $L^*$ ,  $a^*$  and  $b^*$  values of the sample was recorded [4].

#### **3.3.2 Water activity**

Homogenised sample was placed into a plastic Aw tray and flattened with a spatula so that the sample covered the bottom of the tray without exceeding more than half the depth of the tray; this avoided fouling the instrument.

The sample was then placed in the instrument (Aqualab CX-2 water activity meter) and measured until a consistent value was obtained. A duplicate was measured and the average of the two values was recorded [4,5,6].

#### **3.3.3 Titratable Acidity**

A 20 ml aliquot was taken from the filtered solution prepared earlier using a 20 ml volumetric pipette, and added to a 250 ml conical flask. This solution was diluted to 100ml using Mili-Q water (Milipore MiliQ Plus), and to this was added 1ml of 2% Phenolphthalein Indicator (Ajax Chemicals (laboratory chemical grade)). The resultant solution was then titrated using 0.02 M sodium hydroxide solution (pellets, extra pure, Merck, in Milipore MiliQ Plus

water), standardised using a primary standard. This was repeated until at least two concordant results were obtained. The obtained results were averaged and mathematically converted to show how much 0.1M sodium hydroxide solution would be needed to titrate 100 grams of the sample analysed [7].

### **3.3.4 Total Nitrogen**

Pre-weighed samples were placed in Kjeldahl digestion tubes with two copper sulphate catalyst Kjel-tabs (Foss), some quartz anti-refluxing granules (~100mg, BDH), and 25ml concentrated (MERCK 98%, 18 M) sulphuric acid. At least two blank tubes were present per batch. A blank tube consisted of a filter paper identical to that used for the samples, anti-bumping granules, two Kjeltabs, and 25ml concentrated sulphuric acid. The blank, therefore, simulates a sample tube with all the inherent sources of variability, except the sample. The prepared samples were then digested at 420<sup>0</sup> C for 90 minutes under reduced pressure, applied to the samples by an exhaust manifold connected to a specifically designed vacuum pump and fume trap digester unit (Foss Kjelttech 2020), until a green solution was obtained. On cooling, the solution would typically change to a light transparent blue colour, sometimes crystallising into fine needle-like crystals or precipitating as a white amorphous solid. The samples were removed from the heating block, and allowed to cool to room temperature, before being diluted with 70 ml MiliQ water (Millipore MiliQ Plus). If the samples had solidified during cooling, they were redissolved with agitation by either hand or a large sample vortex generator, and gentle heating if required [2].

The digestion process converts any protein or other nitrogen containing compounds present into ammonium sulphate. The distillation process converts the ammonium ions into ammonia, by deprotonation with an excess of sodium hydroxide solution and transferred the ammonia with steam, to a boric acid solution, where the ammonia (a relatively strong Lewis base) chelated to Boric acid, a Lewis acid. This transfer is done in the Kjeldahl distillation unit (Kjeltech 2100). In the distillation unit, 80 ml concentrated (% 40) sodium hydroxide (pellets, extra pure, Merck, in Milipore MiliQ Plus water) was injected into a sample digestion tube containing the diluted digestion product, this neutralised sulphuric acid still present and deprotonates any ammonium ions. Steam was vigorously passed through the solution, then bubbled through a conical flask containing 25ml of 4% boric acid solution (BDH, AnalR grade, in Milipore MiliQ Plus water), to which bromothymol blue methyl red indicator (Sigma-Aldrich) has been added. This was done for 5 minutes. The resultant solution, now blue-green in colour, was titrated using 0.1 M Hydrochloric acid (32% Merck, in Milipore MiliQ Plus water), standardised using a primary standard, until the solution began to change to a red end point (passing through a transparent grey solution just prior to the end point). The volume was recorded, and was mathematically converted to indicate the total nitrogen (or if desired, crude protein) content present in the sample [2,4].

### **3.3.5 Total Lipids determination**

To the pre-prepared 50 ml falcon tube containing the sample, 32 ml of 1% Clarase 40,000 digestive enzyme (Enzyme Solutions, Pty. Ltd.) in 0.5M aqueous sodium acetate solution was added (Merck AnalR, in MiliQ water). The sample was first allowed to hydrate for one hour; the samples were then gently shaken to mix the sample thoroughly with the enzyme solution. The prepared samples were then placed into a water bath heated to 50<sup>0</sup>C and allowed to digest for one hour. The sample was then transferred quantitatively with 80 ml of methanol (Merck, LiChromSol grade) and 40 ml chloroform (BDH, HPLC grade) to a 500ml centrifuge tube where it was thoroughly blended for two minutes with a homogeniser. 40 ml chloroform (BDH, HPLC grade) was added to the centrifuge tube and then the mixture was blended for thirty seconds. 40 ml miliQ water (Milipore MiliQ Plus) was added and following this, was blended for a further 30 seconds. The sample was then centrifuged for 10 minutes at 3000 rpm in a Sorvall refrigerated large volume centrifuge (Sorvall RC-28S), to clarify the chloroform phase. A 20 ml aliquot of the chloroform layer was taken, and transferred to a tared 100ml beaker, an aluminium foil cover was placed on the beaker and it was allowed to stand overnight in a fume hood, to allow the chloroform to evaporate. The samples were then dried to a constant weight in a 101<sup>0</sup> C oven. Next, the samples were allowed to cool in a desiccator to constant weight and then accurately weighed [3].

### **3.3.6 NIRS Analysis**

Samples were scanned in a Foss 6500 Scanning NIR Spectrometer, using a quartz coarse sample cell. The sample cell was two thirds filled with processed fruit packed tightly, ensuring that any voids were plugged with fruit. The spectra were collected in reflectance mode, from 400-2500 nm, using 64 scans. The cell was cleaned using lint free scientific tissues and deionised water. Following this, the cell was left to dry for a few minutes. Between samples, while the cell was drying, and at the beginning of a NIRS session, a reference scan was collected [2,4].

## **3.4 Part II: NIRS assessment of the maturity of unprocessed fruit**

### **3.4.1 Introduction**

Currently, fruit quality assessment is a subjective technique based on fruit colour uniformity and relative lightness of hue, with such limited requirements it is possible to have high quality fruit that fails to meet the expectations of customers. Browning potential is one method of determining a more accurate picture of fruit quality; another potential way of investigation is berry maturity. Such grading is already undertaken in the dried fruit industry, in the United States, where natural sultanas, called “raisins” are sun dried to a blue black colour thus quality assessment by colour is impossible, hence the fruit are graded using an Airstream sorter which separates fruit by their relative density [8,9].

Work by Uhlig *et al*/ has shown a strong relationship between berry maturity by airstream sorter and titratable acidity, and so this technique was used to give a quality parameter strongly linked to maturity, in order to develop a NIRS based calibration [10].

### **3.5 Sampling Procedure**

Samples were initially collected at weekly intervals from the same plants over a period of 6 weeks, from an experimental plot at CSIRO, Merbien. These plants had previously been treated with drying agent then dried. Two kilograms of sample were prepared in this way over the course of the experiment. One kilogram was set aside as a unsorted sample, and the other kilogram was separated into a series of fractions using an airstream sorter. This procedure is used by the Californian natural sultana industry to grade their dark fruit, which is graded on full bodied-ness, not colour.

The recovered fractions were then weighed and kept separate as sorted samples. The rest of the experiment was made up of samples of unprocessed dried fruit collected from a variety of sources throughout the Sunraysia region. This experiment was made up of 134 samples, of these 63 samples were from the experimental plot at CSIRO, Merbien, seven original unsorted samples from each of the harvesting times, and a sample of fruit collected at each airstream sorter setting (in the case of 0.78, two samples) or nine samples per harvest date. A total of 71 unsorted samples were used to ensure that the calibration would be robust.

## **3.6 Sample Preparation**

### **3.6.1 General technique**

Samples were analysed in batches to minimise the time between analysing the fruit with NIRS, and the other techniques used during the experiment. Typically, the samples that were small in size would be analysed with non-invasive techniques first, to ensure that there would be sufficient sample to analyse. The time between sample analysis by NIRS and the other techniques was no more than three days. The samples were cleared of vine trash and stalk material, empty fruit husks and leaf matter prior to any of the analyses that were undertaken as a part of this experiment, but the fruit still contained their cap stems.

### **3.6.2 Sample Preparation for destructive analytical techniques**

Refer to sections 3.2.2 and 3.2.3.

### **3.6.3 Airstream Sorting**

Airstream Sorting was carried out externally at CSIRO, Merbein. For details refer to Appendix 5.

## **3.7 Experimental analysis**

### **3.7.1 CIE Tricolour stimulus**

Refer to section 3.3.1

### **3.7.2 Water Activity by Dew Point**

Refer to section 3.3.2.

### **3.7.3 Titratable Acidity**

Refer to section 3.3.3

### **3.7.4 NIRS Analysis**

The unprocessed samples cleared of vine trash and stem matter were scanned in a Foss 6500 Scanning NIR Spectrometer. Due to the small size of many of the samples of this calibration, it was decided to use, in as many cases as possible, a quarter full coarse sample cell, to ensure that as many samples as possible were large enough to be considered representative during NIRS analysis. In some cases, however, the powder cell was required to measure some of the smaller samples, these samples were labelled with an asterisk (\*) in their sample name so they could be easily identified, if they proved to be detrimental to the calibrations.

For further details of NIRS techniques common to both experiments, refer to section 3.3.6

### **3.8 Calibration Development**

Following collection of NIRS data using NSAS, the spectroscopic data was transferred to Vision (Version 2.22 NIRSystems Proprietary Software) and the laboratory data collected for each sample entered. The calibration was first examined using a Histogram, to determine if the calibration has a “normal” or a “box-car” distribution. If the calibration was determined to have a “normal” distribution, samples within one standard deviation of the mean had half their number randomly removed from the calibration, with the criteria that it made the calibrations distribution more even [11]. The calibration was then checked for spectral outliers using Mahanobis distance in principal component space selection. The sample set was divided into a calibration and validation set, by randomly selecting a fifth of the total set and setting it aside, however it must meet the basic criteria of having a range similar to the original set with a broad distribution. The calibrations were then pre-treated using n-point smooth and 2<sup>nd</sup> derivative pre-treatments prior to calibration development.

### **3.9 Calibration optimisation and outlier removal**

In the case of MLR, calibrations were developed using 6 summation terms, as this number of terms was considered to be acceptable considering the size and nature of the calibrations being developed. An outlier removal limit set to 10% of total samples in the calibration set. The program selects wavelengths that correspond to terms used to develop the calibration. The wavelengths selected as a part of the calibration, often change subtly during calibration development as outliers are removed and a more linear model is developed. PLS Calibrations were developed using factors up to 10% of the number of samples in the calibration set. The number of outliers removed was also limited to a maximum of 10% of the total of samples present. The general approach of outlier removal was common to both MLR and PLS calibration development. The process of outlier removal is a gradual and methodical process, which starts with the examination of the calibration plot, the residuals versus predicted values and the residuals versus laboratory data plots to look for the outermost sample points. A short list of potential outlier points was recorded, and each one is removed from the plot, the calibration is then redeveloped in its absence and the resultant Correlation and SEC is recorded, the point that was removed is returned to the calibration, and the next point in the short list is removed and the process continues until all the potential outliers are tried. The point with the most detrimental effect on the calibration is removed and the calibration is optimised, and a new short list for the next outlier is made. The above procedure was continued until the calibration gave a correlation of greater than 0.85 or 10% of the calibration set had been removed as outliers. This limit was chosen due to the nature of the subject

being studied by NIRS, and limits the possibility of the calibration being over-fitted and unrepresentative of “real world” samples, as it is preferable to remove as few outliers as possible to achieve the desired outcome.

The resultant calibrations that have been developed were then compared to the validation set, a set of data points set aside which was of similar range and distribution to the calibration set, with which the developed calibration is tested. The predictions of the analyte of interest based on the spectroscopic data were plotted against the physicochemical data, and the resultant correlation of this plot was obtained as an unbiased measure of how well the developed calibration would predict in the packing shed.

### 3.10 Bibliography

1. AOAC Official Method 920.149
2. *Vogel's Textbook of Quantitative Chemical Analysis*, Fifth Edition. New York: Longman Scientific and Technical; 1989.
3. AOAC Official Method 983.23
4. Frank, D. *Investigation of the biochemical basis of browning during the storage of sultanas, a PhD Thesis*. Victoria University, Werribee Campus; 2001.
5. Huxoll, CC. "Near infrared analysis potential for measuring the water activity of raisins". *Conference Proceedings of the IFT Annual Meeting* 1995: p.70
6. Huxoll, CC. "Assessment of Near Infrared (NIR) Diffuse Reflectance analysis for measuring Moisture and Water Activity in Raisins", *Journal of Food Processing and Preservation* 2000;24(4): 315-333.
7. AOAC Official Method 942.15
8. AOAC 6<sup>th</sup> edition. p596. 1945
9. Kasimatis, AN, Vilas, EP Jr., Swanson, FH, Baranek, PB. "Relationship of Soluble Solids and Berry Weight to Airstream Grades", *American Journal of Enology and Viticulture* 1977; 28(1): 8-15.
10. Uhlig, BA, Clingeleffer, PR. "Influence of Grape (*Vitis Vinifera* L.) Berry Maturity on Dried Fruit Colour" *Journal of Horticultural Science and Biotechnology* 1998; 73(3): 329-339.
11. Brimmer, PJ, Hall, JW. "Method development and implementation of near infrared spectroscopy in industrial manufacturing support laboratories." In: Williams, PC, Norris, K, editors. *Near Infrared Technology in the Agricultural and Food Industries*, 2<sup>nd</sup> Edition. St. Paul (Mi): American Association of Cereal Chemists; 2001. p. 193-194

## **4.0 Results and Discussion I: Quality Assurance of Processed Sultanas – CIE L\* a\* and b\* Values**

### **4.1 Introduction**

Previous work has enabled a number of key quality parameters in dried fruit to be analysed prior to processing, at the transaction between grower and processor, with the aim of improving the objectivity of quality assessment, ensuring a truer assessment of fruit quality and storage stability [1]. A logical next step was the application of similar assessment techniques to the point of transaction between processor and buyer, with the aim to enable the processor to meet the needs of the customer with more precision. These aims would be more readily met by providing the processor with an objective means of rapidly assessing colour, active water and nitrogen content, which may have changed during the time that the fruit remained in storage. This information in turn would allow the processor to make blends of fruit to meet the customers' requirements. These blends, however, no longer completely conform to the conventions of crown grades. The crown grade assessment occurs during the grower-processor transfer, and is often not as well defined by the time the fruit is sold to customers.

Experimental hypothesis:

- It would be feasible to develop calibrations capable of determining colour of processed fruit, thus allowing objective assessment of colour to more readily meet customer requirements.

## 4.2 Objectives

The primary aim of the following research was to collect samples over three seasons, from the processors in the Sunraysia Region, of as wide a variety of fruit crown grades and maturities as possible. Analyse these series of samples with CIE Tricolour Stimulus using a Minolta Chromameter. Develop robust and proven calibrations using both MLR and PLS regression analysis techniques, using NIRS diffuse reflectance spectroscopy. Once calibrations have been developed, trial NIRS as a technique of rapid assessment of processed samples.

## 4.3 Inter-correlation analysis

### 4.3.1 Year 1

Below is a correlation matrix (Table 4.1), which presents the level of similarity between of a series of observations, in this case chemical and physicochemical analyses made on a series of samples from the first season of the project.

Table 4.1 Inter-correlations between constituents of the first seasons' laboratory data

Constituents	CIE L*	CIE a*	CIE b*	Aw	Protein	Titrateable Acidity
CIE L*	1.0000	-0.0184	0.5933	0.3889	-0.2123	-0.2831
CIE a*	-0.0184	1.0000	-0.3079	-0.0343	0.3460	-0.4209
CIE b*	<b>0.5933</b>	-0.3079	1.0000	0.3976	-0.1234	-0.2343
Aw	0.3889	-0.0343	0.3976	1.0000	-0.1882	-0.1702
Protein	-0.2123	0.3460	-0.1234	-0.1882	1.0000	0.0543
Titrateable Acidity	-0.2831	<b>-0.4209</b>	-0.2343	-0.1702	0.0543	1.0000

The only strong relationship seen within the range of data were examined in this seasons calibrations, was between CIE L\* and CIE b\* (Table 4.1, CIE L\* versus CIE b\*=0.59), surprisingly there was no strong link between CIE a\* and

L\* or b\*. The next strongest association was CIE a\* and titratable acidity, (Table 4.1 CIE a\* versus Titratable acidity= -0.42), as there were established links between colour and maturity, of which titratable acidity was an indirect indication, this was not unexpected. These values, however, were preliminary as they were based on one season's fruit.

#### 4.3.2 Year 2

In the correlations between constituents of the second season showed similar trends to those observed in the previous season. CIE L\* showed a strong relationship with CIE b\* (Table 4.2, CIE L\* versus CIE b\*=0.78), and a notable inverse relationship with CIE a\* (Table 4.2, CIE L\* versus CIE a\*= -0.42). CIE a\*, as well as those already noted showed a strong inverse relationship with CIE b\* (Table 4.2, CIE a\* versus CIE b\*= -0.735).

Table 4.2 Matrix of correlations between constituents of the second seasons' laboratory data

Constituents	CIE L*	CIE a*	CIE b*	Aw	Protein	Titratable Acidity
CIE L*	1.0000	-0.4237	0.7781	0.0352	-0.3339	0.0927
CIE a*	<b>-0.4237</b>	1.0000	-0.7350	-0.1777	-0.0754	-0.3126
CIE b*	<b>0.7781</b>	<b>-0.7350</b>	1.0000	-0.0091	-0.2466	0.2389
Aw	0.0352	-0.1777	-0.0091	1.0000	0.0698	0.1893
Protein	-0.3339	-0.0754	-0.2466	0.0698	1.0000	0.4747
Titratable Acidity	0.0927	-0.3126	0.2389	0.1893	<b>0.4747</b>	1.0000

Kjeldahl Protein exhibited an interesting relationship with titratable acidity (Table 4.2, Kjeldahl Protein versus titratable acidity= -0.475). While the data in this comparison was more defined than the previous season it was still preliminary and should be verified across three seasons to establish if the trend is continued and not due to extreme climatic conditions in one of the seasons.

### 4.3.3 Year 3

In contrast to the previous season, only a few inter-constituent relationships were seen in the third season, however one of these relationships was seen previously. CIE L\* had a notable association with CIE b\* (Table 4.3 CIE L\* versus CIE b\*= 0.46), much less pronounced than in previous seasons, but present nevertheless.

Table 4.3 Matrix of correlations between constituents of the third seasons' laboratory data

Constituents	CIE L*	CIE a*	CIE b*	Aw	Protein	Titrateable Acidity
CIE L*	1.0000	0.0692	0.4623	0.1417	0.0841	-0.0498
CIE a*	0.0692	1.0000	-0.2719	0.1796	-0.4632	-0.0619
CIE b*	<b>0.4623</b>	-0.2719	1.0000	-0.1912	0.0366	0.1343
Aw	0.1417	0.1796	-0.1912	1.0000	-0.0477	-0.1004
Protein	0.0841	<b>-0.4632</b>	0.0366	-0.0477	1.0000	0.0630
Titrateable Acidity	-0.0498	-0.0619	0.1343	-0.1004	0.0630	1.0000

CIE a\* showed a weak inverse relationship with Kjeldahl Protein (Table 4.3 CIE a\* versus KP= -0.46), previously a link has been present between these two components, but the relationship was reversed this season. This could be due to variations in climate, during the growth phase and when the samples were on the drying rack.

#### 4.3.4 Final Combined Calibration

The final comparison between constituents across all three seasons gave a true indication of associations between constituents, removing any false trends due to seasonal variation and a size that can give a definite trend in results.

Table 4.4 Matrix of correlations between constituents of the combined seasons' laboratory data

Constituents	CIE L*	CIE a*	CIE b*	Aw	Protein	Titrateable Acidity
CIE L*	1.0000	-0.1757	0.6582	0.1162	-0.1565	0.0335
CIE a*	-0.1757	1.0000	-0.4513	0.0618	-0.2036	-0.1694
CIE b*	<b>0.6582</b>	<b>-0.4513</b>	1.0000	-0.0211	-0.1190	0.1763
Aw	0.1162	0.0618	-0.0211	1.0000	-0.0629	0.0061
Protein	-0.1565	-0.2036	-0.1190	-0.0629	1.0000	0.1930
Titrateable Acidity	0.0335	-0.1694	0.1763	0.0061	0.1930	1.0000

CIE L\* consistently showed an association between itself and CIE b\* (Table 4.4 CIE L\* versus CIE b\*= 0.66), and this was shown again in the combined comparison. CIE a\* showed a weak inverse relationship between itself and CIE b\* (Table CIE a\* versus CIE b\*=-0.45), which was also seen in individual seasons of the project.

#### 4.4 Calibration development technique

The development of a calibration using MLR with a sample size greater than 150, sufficient for use in an industrial application, was developed using 6 summation terms [2]. Normally this would be considered over-fitting, however due to the increased variance of a larger calibration, the greater number of terms were required to take account of this variance [2]. This use of an increased number of summation terms has been often used when the medium being analysed is non-homogenous, as is the case here, or in a calibration that only showed a weak association of the analyte of interest to the NIR spectra. Pre-treatments for the calibrations in all cases were N-point smooth and 2<sup>nd</sup> derivative. In the case of PLS calibrations, similar to MLR the amount of mathematical compensation used to develop a calibration was limited by its size. The number of factors used to optimise the calibration was set at 10% of the total number of samples within the calibration set, however, if an acceptable performing calibration could be developed using less factors without undue sacrifice of correlation or calibration error, this was done [3]. As stated previously, an upper limit of 10% of the samples in the calibration set for removal of outliers is fixed. Ideally, there should be no need to remove outliers at all from the calibration set, as the removal of each sample means there would be less variance explained by the calibration. This is not always possible, as the nature of the medium being analysed, whole dried sultanas, are not homogenous or sufficiently small in size to become homogenous with mixing. During calibration development, therefore, it was necessary to remove outliers from the calibration set.

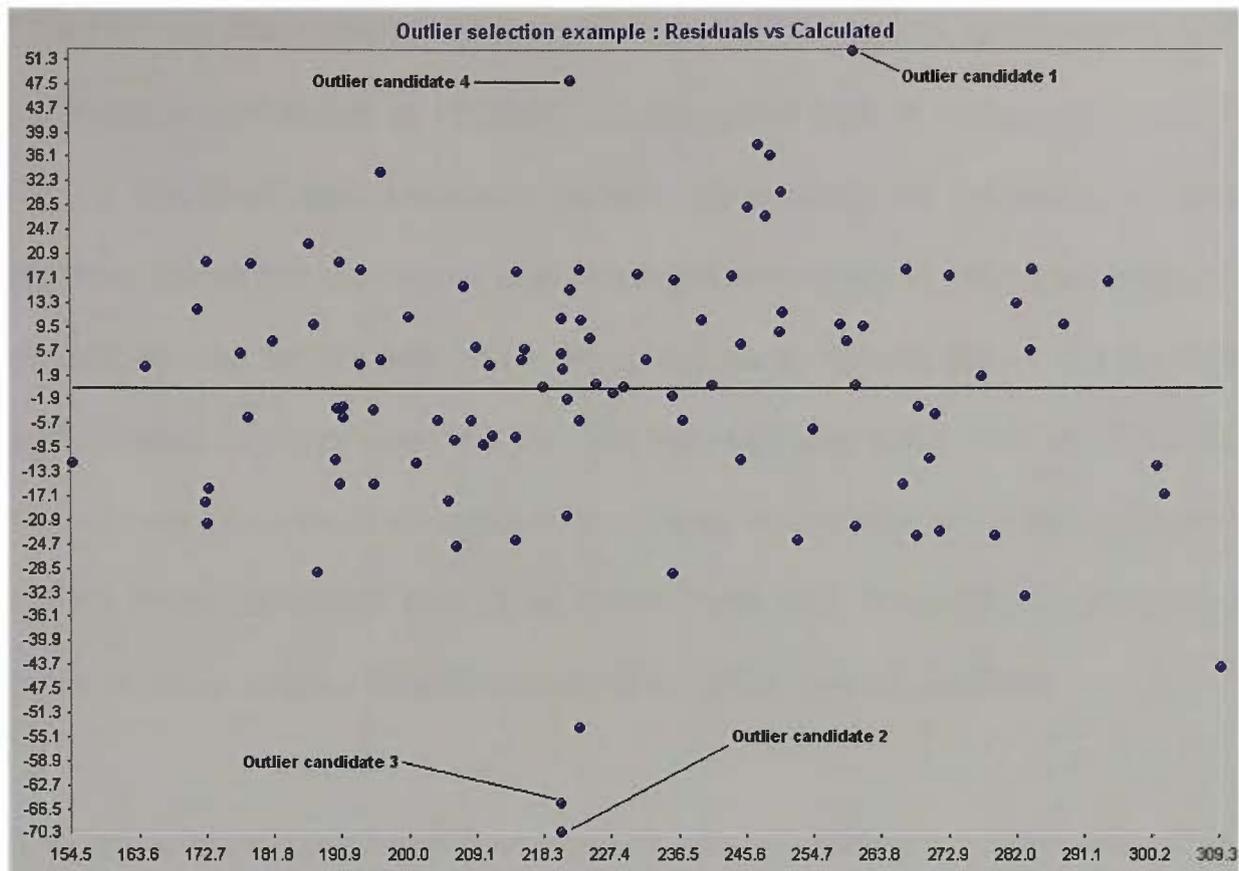


Figure 4.1 A residual plot illustrating outlier selection.

Outlier selection during calibration development is the most time consuming part of the calibration, the calibration is viewed as a residual plot versus either the laboratory values or the calculated values (Figure 4.1). This plot is then examined and a number of samples are chosen as likely outlier candidates, these are typically chosen by their relative distance from the line of best fit that passes through the samples. These samples were identified, recorded then each sample was removed from the calibration individually and the calibration was then recalculated.

The result for each sample was recorded, and the sample that showed a clear improvement in the calibration was used, in this way a sample can be easily returned to the calibration if the point appeared to be a leverage point that was important in the calibrations performance, rather than as an outlier [4].

This process was repeated until a calibration of acceptable performance was achieved (a correlation of  $R^2 > 0.85$ ), or the upper limit of outliers is reached. [5]. If this level was exceeded without the removal of outliers, or it was reached before the upper limit of outliers was removed the calibration was not developed any further and was considered ready for use [6]. If this level of performance was not reached, but fell between this value and an  $R^2 = 0.75$ , then it was considered acceptable, and ready for testing using the validation set [5]. If the calibration was close to the lower limit of correlation, it may be possible to re-develop the calibration, altering the outliers removed.

This approach to calibration development was used throughout this project for both sets of calibrations, for the processed fruit experiment and the maturity assessment experiment.

## 4.5 Chemical assignments and NIRS

It has always been considered important to link the calibration developed from NIR data to the chemical constituent of interest, however, considering the diverse nature of compounds present in natural produce and the subtle interactions occurring between compounds within an agricultural project, it is impossible to rule out a relationship between seemingly unrelated compounds or functional groups without further investigation.

### Coloured compounds

Coloured compounds are primarily of two types. Highly conjugated pi systems within organic compounds with or without heteroatoms such as nitrogen or oxygen, known as chromophores. The others are transition metal compounds, either inorganic or coordinated to organic chelates [7].

As more double bonds are included in a pi system, the wavelength of light that is absorbed lengthens resulting in a range of coloured compounds from colourless through yellow to red [8]. Similarly the presence of heteroatoms or the addition of groups containing heteroatoms also increases the wavelength of light that is absorbed [7].

Prime examples of conjugated coloured compounds containing heteroatoms within sultanas are the products of the glucose-arginine Maillard reaction, some among many compounds caused by this reaction are 5-hydroxymethylfurfural and 4-hydroxymethyl imidazole [1].

Compounds containing transition metal complexes are also cause the absorbance of light within the visible region of the spectrum, such as Chlorophyll a and b [7]. The variety of compounds that are associated with

colour clearly illustrate the origins of wavelengths associated with organic functional groups that contribute to CIE tricolour stimulus calibrations, since many compounds directly influence sultana colour.

#### Amino acids and proteins

A wide variety of organic moieties are associated with the 20 commonly occurring amino acids and their derivatives as well as the proteins they form.

As well as carboxylic acids and amines, there are methyl, isopropyl, *sec*-butyl, benzyl, *p*-hydroxy benzyl groups, amides, primary and secondary hydroxy groups, thiols and sulphides, imines and indoles present within amino acids, the building blocks of proteins, not to mention the modifications to amino acids once proteins are formed [9].

#### Water activity

Often calibrations developed for a variety of constituents give wavelengths linked to water, because water is often present as a hydration sphere around compounds of interest, this is known as the matrix effect, and shows how a seemingly unrelated source of information can be related to the analyte of interest [10] Conversely, compounds associated with water can contribute to a calibration of water activity.

#### Contributing compounds to total acidity

There are a number of compounds present within sultanas, which contribute to their total acidity. Primarily there is malic and tartaric acid, both are n-butyl-1,4-dicarboxylic acids, Malic has a single hydroxy group on the “2” position, while tartaric acid has two hydroxy groups on the “2” and “3” positions. As

well, salicylic (*o*-hydroxybenzoic acid) and *p*-hydroxybenzoic acid are also present. Other common acids within fruit are *p*-coumaric, ferulic, caffeic and sinapic acids. It should also be remembered that any acidic compound contributes to the total acidity, which would therefore include Lewis acids such as Chlorophyll, and the acidic amino acids aspartate and glutamate [11].

## 4.6 Tricolour Stimulus L\*

### 4.6.1 Season one CIE L\*

Table 4.5 Laboratory parameters summary of season one CIE L\*

# Samples (in calibration)	# Samples in validation	Range	Mean	Standard Deviation
111 (89)	22	37.17-27.23	33.12	1.75

In the table above (Table 4.5) is presented the laboratory data of the first seasons calibration for CIE L\* is shown. The number of samples used in this calibration was 111, where 89 of these were used in the training calibration set, prior to the removal of outliers, the remaining 22 samples were set aside as a representative validation set to confirm the validity of the calibration. The samples, CIE L\* values, ranged from 27.23 to 37.17 with a mean of 33.12 and standard deviation of 1.75.

#### 4.6.1.1 MLR calibration

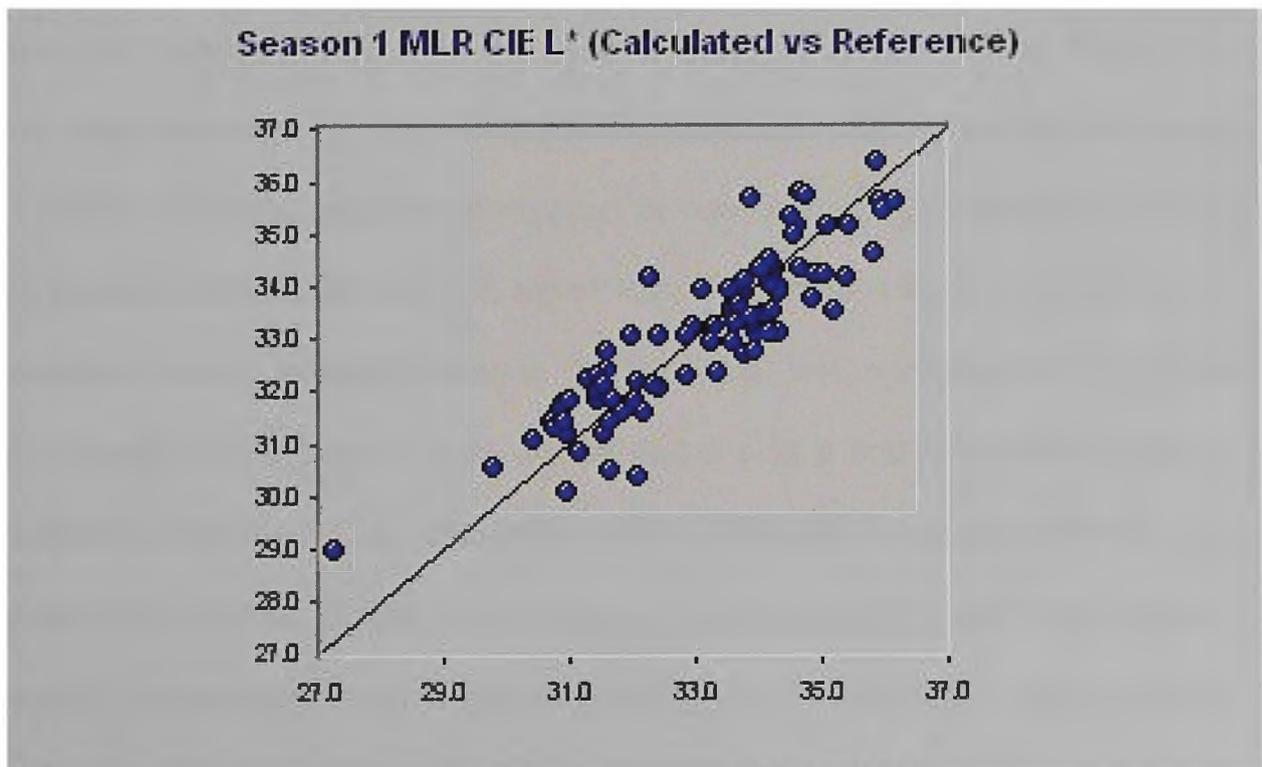


Figure 4.2 A plot of the first seasons calibration for CIE L\* using MLR.

The following assignments are theoretical and as such there are a number of possible origins consistent with the complex matrix of the medium being studied. Further investigation beyond the scope of this project would need to be completed to confirm these assignments. Refer to section 4.6 for further details. The assignments are cross-referenced from three sources, the NIRSystems Chart *Near Infrared Adsorptions*[12], *Near Infrared technology in the Agricultural and Food Industries* edited by Phil Williams and Karl Norris [13], and *Near Infrared Spectroscopy in Food Analysis* by Brian Osborne, Tom Fearn and Peter Hindle[14].

The assignments for the calibration developed for the first year CIE L\* values (Table 4.6), did not show any direct link to the visible region, but the terms used originated from regions that were attributed to the chemical composition of the fruit. The first term (914nm) was attributable to a carbon-hydrogen bond stretch of alkyl groups and alkenes with three organic moieties attached and one hydrogen within the fruit [13], the wavelength indicating that it was the third overtone of such a stretching vibration [12]. The second term (1784nm) was of a wavelength typical of carbon–hydrogen bond-stretching, particularly within cellulose [13], typical functional groups such as alkanes and alkenes except terminal alkenes ( $=CH_2$ ) and aryl compounds [13]. The wavelength of the second term shows that it was a first overtone of such a stretching vibration [12]. The third term (1894nm) was attributed to the stretching vibration of the carbon-oxygen double bond in carbonyl groups, typically in carboxylic acid functional groups [14] or even carbon oxygen single bonds in secondary, tertiary or aryl hydroxy groups, or water [13].

The wavelength (2354nm) of the fourth term, from the combinations band [12], was attributed to the deformation vibration of a carbon-hydrogen bond, typically attributed to cellulose [14], such a feature was due to the second overtone band of such a deformation [12].

Table 4.6 The wavelengths of the first seasons MLR calibration of CIE L\* (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

MLR					
Term	Outliers	$\lambda$ (nm)	R <sup>2</sup> (Cal)	SEC	F-value
1	0	914	0.517	1.187	87.8
2	“	1784	0.635	1.038	70.52
3	“	1894	0.715	0.922	67.04
<b>4</b>	<b>0</b>	<b>2354</b>	<b>0.749</b>	<b>0.873</b>	<b>58.8</b>

A plot of the MLR Calibration set is included above (Figure 4.2) as well as a table displaying the wavelengths used in this calibration (Table 4.6). The resultant calibration of CIE L\* was developed by measuring 111 processed samples with a Minolta Chromameter, using the CIE colour system. The range and distribution (Table 4.5) of the samples collected over the season, was diverse as L\* is a critical parameter in influencing the crown grade of fruit. This parameter was primarily influenced by the relative lightness of the fruit, and with a broad spread of crown grades assessed in resulted in a relatively broad range of values. The resultant calibration was not completely representative, as it was made of samples from only one season. The purpose of these seasonal examinations was to allow for variations in fruit quality due to changes in climatic conditions thus monitoring the progress of the calibration over the length of the project. As expected, the calibration developed using MLR of the first seasons work was limited in its ability to show a trend between the features on NIRS spectra and corresponding laboratory data.

Naturally as more summation terms were added to the calibration, more of the variance within the spectral data was compensated. The higher the correlation coefficient and the smaller the standard error of calibration (Table 4.6 Term=1  $R^2=0.517$ , SEC=1.19; Term=4  $R^2=0.749$ , SEC=0.873), however the more terms or factors in the case of PLS, then the calibration was more suited to explaining the variance in the calibration sample set rather than the ability to predict new samples. This means there must be a balance between improving the calibration and robustness of prediction. As this is a procedure that was followed through out the calibration development, it will not be covered further to any great extent in the remainder of the thesis. The trend seen with the F-value as more summation terms are added to the calibration was a decrease in size (Table 4.7 F-value trend: 87.8 to 58.8). This indicated that the calibration was of value and not over-fitted [15].

#### 4.6.1.2 PLS Calibration of CIE L\*

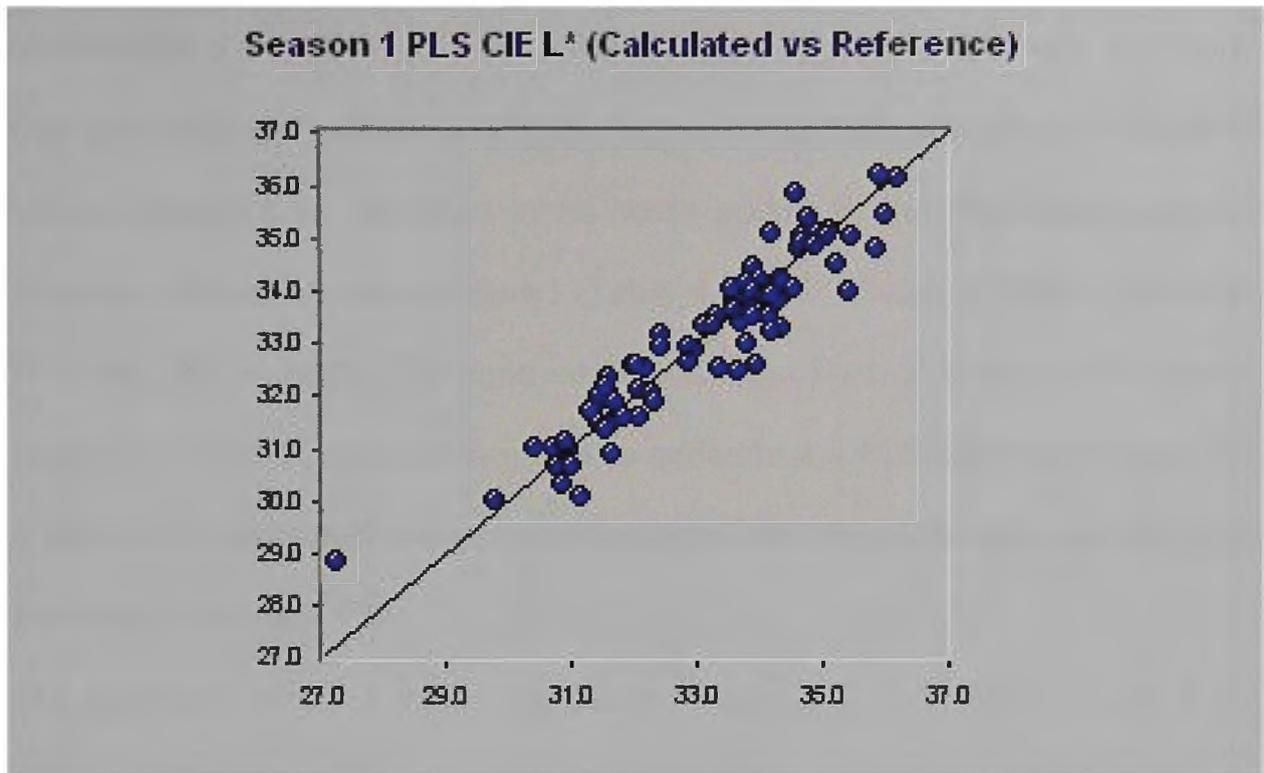


Figure 4.3 A plot of the first seasons calibration for CIE L\* using PLS.

Interestingly, the correlation obtained with PLS (Table 4.7  $R^2=0.89$ ) for the same samples was much improved over MLR (Table 4.6  $R^2=0.75$ ), this was attributed to the fact that PLS is a spectral technique and used trends seen across the entire spectrum, where the improved correlation was then easily explained [16].

Table 4.7 The factors of the first seasons PLS calibration (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

PLS					
Factors	Outliers	$R^2$ (Cal)	SEC	SECV	F-value
1	-	0.372	1.35	1.38	48.56
2	-	0.592	1.10	1.26	58.66
3	-	0.656	1.01	1.20	50.77
4	-	0.685	0.977	1.19	42.88
5	-	0.776	0.829	1.20	54.06
6	-	0.807	0.775	1.18	53.51
7	-	0.837	0.715	1.14	55.77
8	-	0.854	0.681	1.07	54.99
9	-	<b>0.892</b>	<b>0.592</b>	<b>1.02</b>	<b>67.68</b>

It should also be observed, however, that this was a calibration of a single season and a stronger indicator of NIR's ability to predict CIE L\* was obtained over time with more diverse samples. A plot of this PLS calibration is included above (Figure 4.3). As more terms were added to the PLS calibration a stronger correlation was achieved (Table 4.7 F=1  $R^2=0.37$ , SEC=1.35; F=9  $R^2=0.98$ , SEC=0.592). The trend seen within the F-value (Table 4.7 F-value trend: 48.6 to 67.7) as more terms were added to the PLS calibration is one of a slight increase, which was a clear indicator that the calibration was a good estimate of the data [15].

The standard error of cross validation (Table 4.7 F=1 SECV=1.38; F=9 SECV=1.02) showed a clear decrease in values as the final factor was approached during calibration development, this indicated that adding further factors into the calibration would not be detrimental [17].

## 4.6.2 Season two CIE L\*

Table 4.8 Laboratory parameters summary of season two L\*

# Samples (in calibration)	# Samples in validation	Range	Mean	Standard Deviation
95 (76)	19	28.41-36.6	32.52	1.98

### 4.6.2.1 MLR Calibration

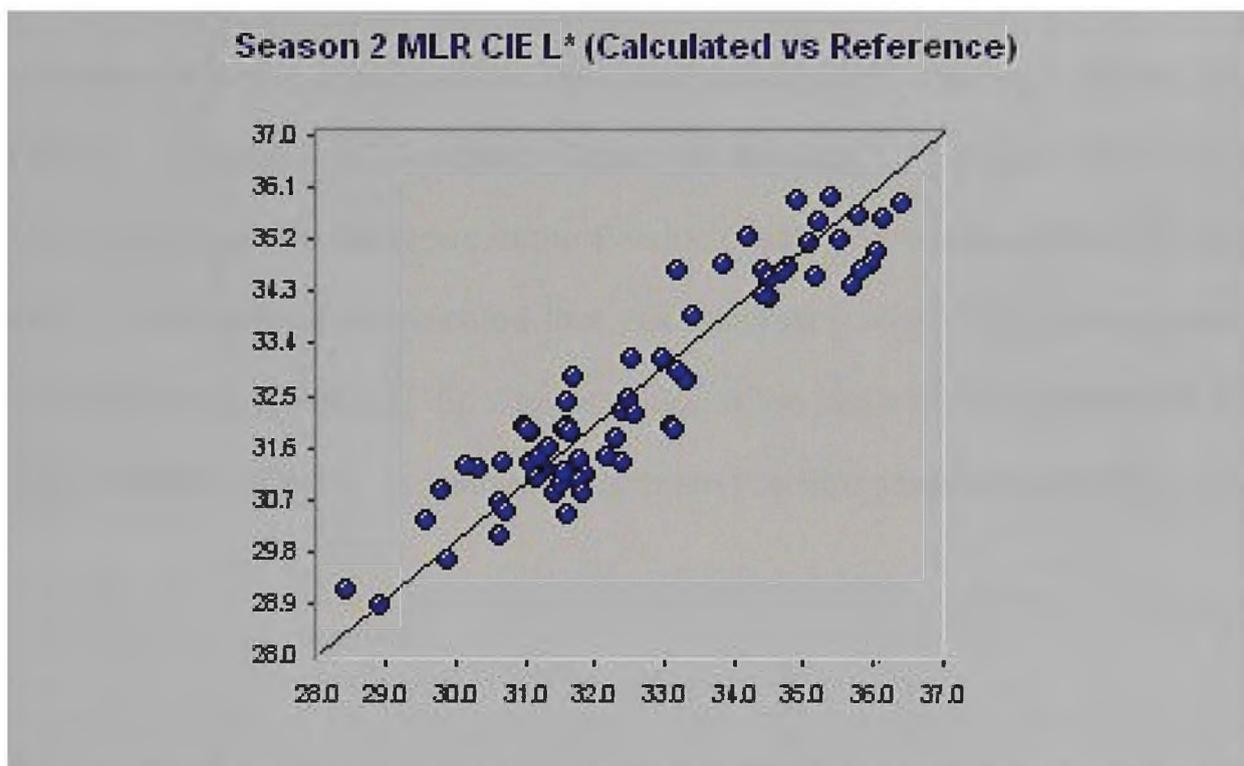


Figure 4.4 A plot of the second seasons calibration for CIE L\* using MLR.

The first three terms (Table 4.9 wavelengths: 514nm, 566nm, 486nm respectively) used to develop this calibration all originated from the visible region of the NIR Spectra [12,13,14], however, the final term (1982nm), from the combinations band [12], was linked to the asymmetric stretching of nitrogen-hydrogen bonds and the vibrations of secondary amides [13], typically associated with protein [14]. This wavelength was also linked to secondary and tertiary hydroxy groups, carboxylic acids, water and phenols, and even terminal alkenes ( $=CH_2$ ) [13]. Compounds such as arginine can be substrates of reactions that cause highly coloured compounds or are part of highly coloured compounds themselves [1].

Table 4.9 The wavelengths contributing to the second seasons MLR calibration (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

MLR					
Term	Outliers	$\lambda$ (nm)	$R^2$ (Cal)	SEC	F-value
1	-	514	0.7599	0.967	215.27
2	-	566	0.807	0.874	139.83
3	-	486	0.830	0.8258	107.47
4	-	<b>1982</b>	<b>0.845</b>	<b>0.794</b>	<b>88.66</b>

The use of more summation terms in the calibration resulted in an improved correlation with the spectroscopic data, and equally SEC improved (Table 4.9 Term=1  $R^2=0.76$ , SEC=0.967; Term 4  $R^2=0.845$ , SEC=0.794). The comparatively sharp decrease in the F-value (Table 4.9 F-value trend: 215 to 88.7) of this calibration indicated that this calibration might have been more susceptible to over-fitting [15]. This was due to the fact that the samples were not sufficiently diverse. In a larger sample set this would be overcome [2].

#### 4.6.2.2 PLS Calibration

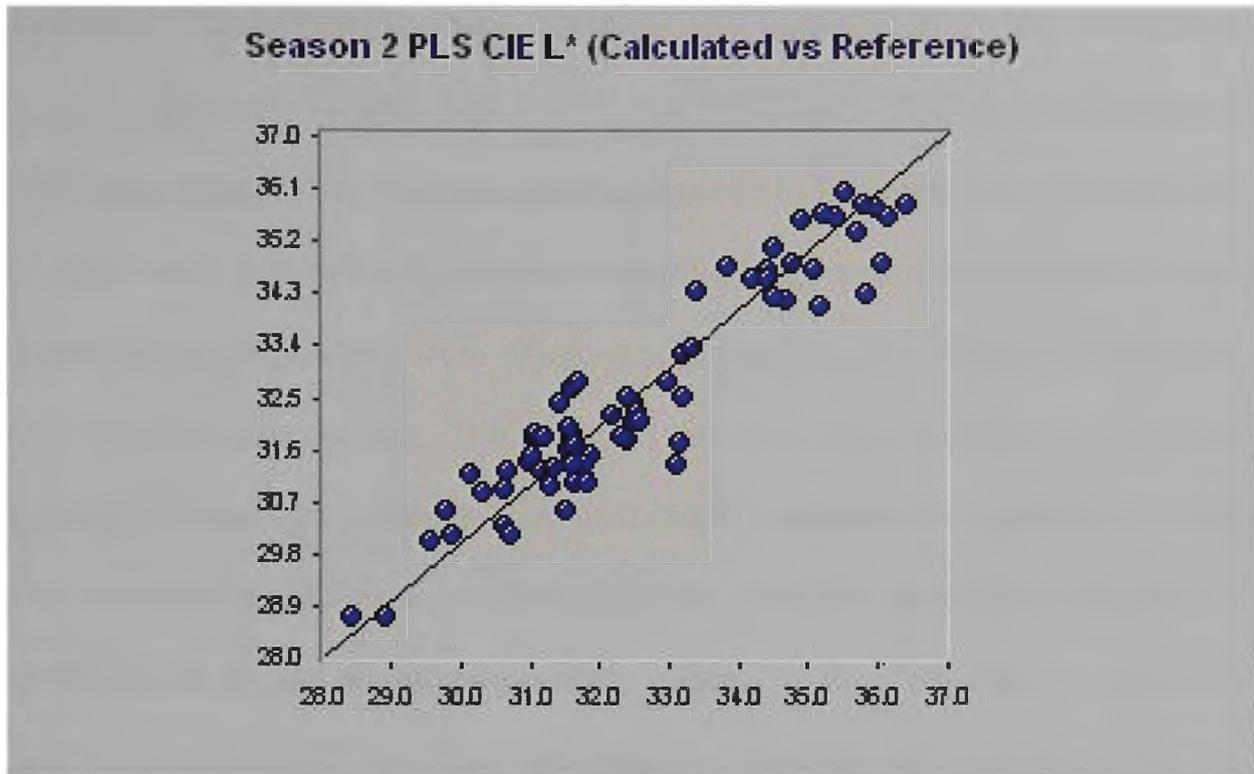


Figure 4.5 A plot of the second seasons calibration for CIE L\* using PLS

In comparing the two calibrations of the second season of the study it was apparent that the PLS (Table 4.10) had a performance edge over MLR (Table 4.9), both its correlation and SEC indicate it was more capable of prediction than its MLR equivalent, however the F-value was lower (Table 4.9 MLR F-value=88.7 compared to table 4.10 PLS F-value=65.3) which indicated that there may be less variation in these samples than those previously encountered. A plot of the calibration has been included above (Figure 4.5).

Table 4.10 The factors contributing to the second seasons PLS calibration of CIE L\* (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

PLS					
Factors	Outliers	R <sup>2</sup> (Cal)	SEC	SECV	F-value
1	-	0.638	1.187	1.205	119.7
2	-	0.717	1.057	1.150	85.03
3	-	0.808	0.877	1.027	92.79
4	-	0.844	0.797	1.036	87.89
5	-	0.850	0.789	0.999	72.37
6	-	0.867	0.747	1.035	68.61
7	-	0.877	0.725	1.047	63.05
<b>8</b>	-	<b>0.895</b>	<b>0.674</b>	<b>1.050</b>	<b>65.34</b>

As more terms were added to the PLS calibration, an improvement in the coefficient of determination was achieved and therefore SEC also improved (Table 4.10 F=1  $R^2=0.64$ , SEC=1.19; F=8  $R^2=0.895$ , SEC=0.674). Again a steady decrease in the F-value was encountered in this calibration (Table 4.10 F-value trend: 120 to 65.3), this confirmed the requirement to analyse a more diverse range of samples from different seasons [15]. The data here however was only from one season. The standard error of cross validation, used to get a gauge of how well the developed calibration fitted to a cross validation set, and therefore predict new samples indicates that this value, has reached a minima and began to increase again (Table 4.10, F=1 SECV=1.2; F=5 SECV=0.99; F=8 SECV=1.05). This indicated that the calibration was slightly over-fitted, however as this value increased only gradually, any over-fitting was minimal [17].

### 4.6.3 Season three CIE L\*

Table 4.11 Laboratory parameters summary of season three L\*

# Samples (in calibration)	# Samples in validation	Range	Mean	Standard Deviation
108 (86)	22	28.76-35.14	31.96	1.31

#### 4.6.3.1 MLR calibration

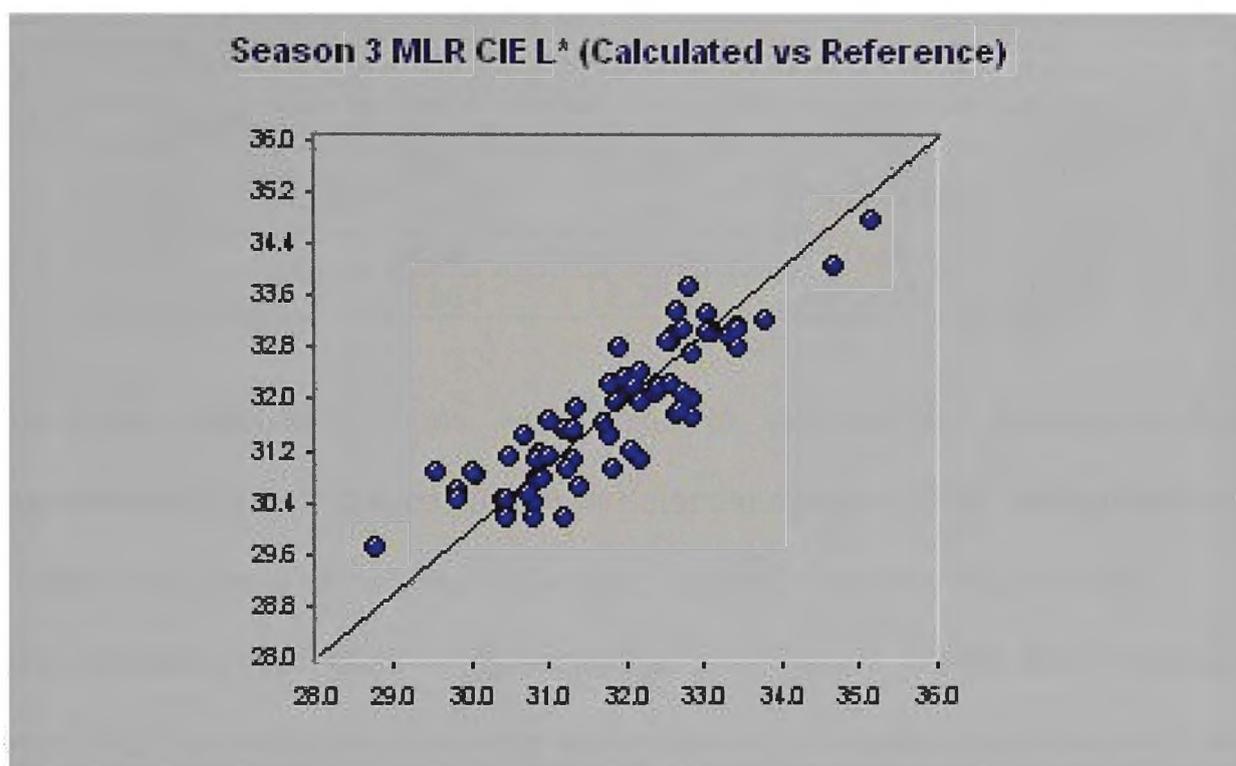


Figure 4.6 A plot of the third seasons calibration for CIE L\* using MLR.

The calibration of the third season of samples based on the CIE parameter L\* (Table 4.12) did not exhibit a strong link to the visible region of the spectra, the features that were more prominent were due to functional group vibrations, these would be linked to strongly coloured compounds that contribute to the overall colour of the fruit. The first term (896nm) of this calibration could be linked to the third overtone stretching vibrations of carbon hydrogen bonds [12] within alkyl groups, aldehydes, and terminal alkenes [13]. The second term (1544nm) was associated with the stretching vibrations of oxygen hydrogen bonds [12] within hydroxy groups, often due to hydrogen bonding within starch [14], as well as amines and water [13]. The third term

(420nm) comes from the visible region of the spectrum [13]. Asymmetric stretching vibrations of nitrogen hydrogen bonds in primary amides were the likely source of the fourth term (1964nm), and aqueous ammonia as well as tertiary, carboxylic acid and aryl hydroxy groups, water and the carbonyl of ketones [13]. Included above is a plot of the MLR calibration (Figure 4.6).

Table 4.12 The wavelengths contributing to the third seasons MLR CIE L\* calibration (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

MLR					
Term	Outliers	$\lambda$ (nm)	R <sup>2</sup> (Cal)	SEC	F-value
1	9	896	0.536	0.836	78.69
2	"	1544	0.677	0.700	70.16
3	"	420	0.778	0.585	76.92
4	9	1964	0.788	0.575	60.5

As more summation terms were used to develop the calibration an improvement in both the coefficient of determination and SEC was achieved (Table 4.12 Term=1 R<sup>2</sup>=0.54, SEC=0.84; Term=4 R<sup>2</sup>=0.788, SEC=0.575).

This calibration exhibited a slight decrease in F-value (Table 4.12 F-value trend: 78.7 to 60.5) as more terms were added to the calibration, however the rate of decrease was only slight [15]. Once the three seasons calibrations were combined a more robust calibration was achieved (see section 4.6.4).

### 4.6.3.2 PLS Calibration

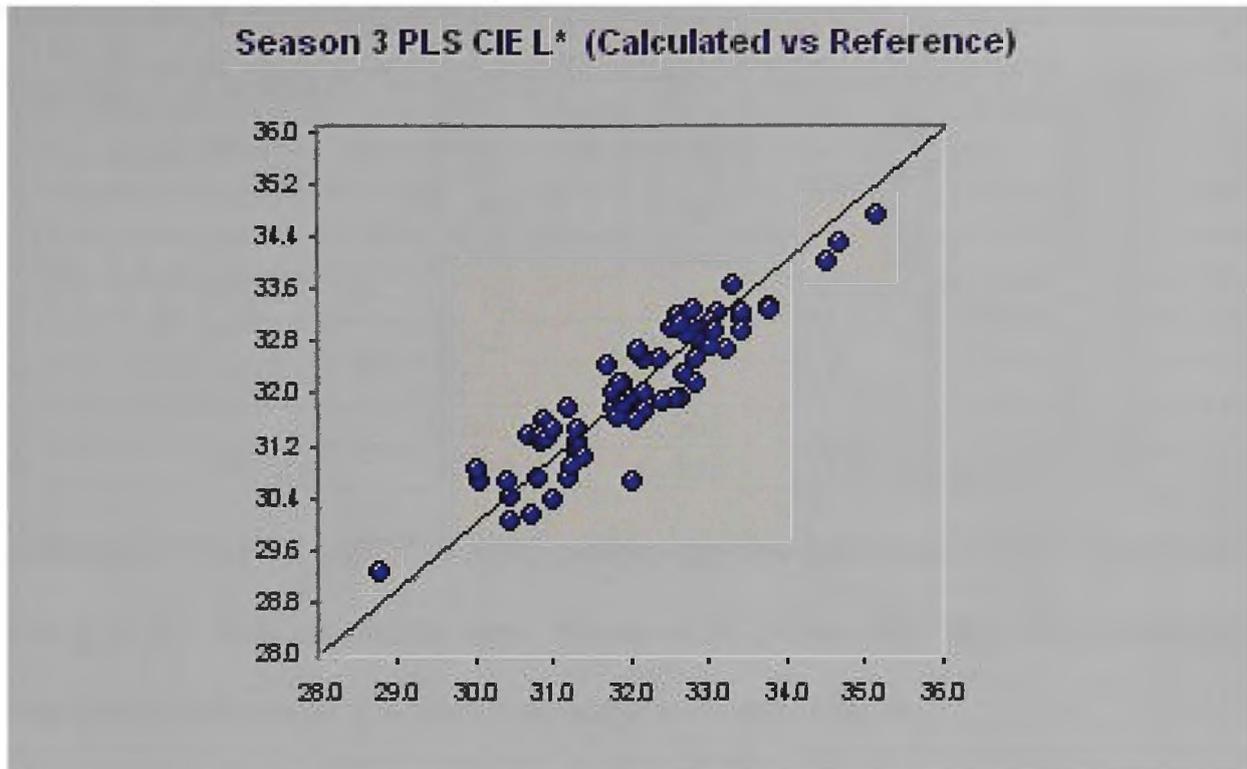


Figure 4.7 A plot of the third seasons calibration for CIE L\* Using PLS

Refer to the above-included plot of this calibration (Figure 4.7). Comparing the calibrations of the third season of results reinforced the general trend that PLS out-performs MLR, which was clearly demonstrated in their respective correlation and SEC values (Table 4.12 MLR  $R^2=0.788$ ,  $SEC=0.575$ ; Table 4.13 PLS  $R^2=0.854$ ,  $SEC=0.486$ ), the PLS calibration had a much lower F-value (Table 4.12 MLR F-value=60.5; Table 4.13 PLS F-value=39.1), this should improve once the calibrations are combined, and a greater sample variety introduced.

As more factors were added to the calibration there was an overall increase in the F-value (Table 4.13 F-value trend: 35.69 to 39.1), this clearly indicated that there was minimal impact in increasing the performance of the calibration [15]. Error of cross validation indicated a trend of initial decrease that reached a minimum then began to increase again (Table 4.13: F=1  $SECV=0.988$ ; F=5

Table 4.13 The factors contributing to the third seasons PLS calibration of CIE L\* (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

PLS					
Factors	Outliers	R <sup>2</sup> (Cal)	SEC	SECV	F-value
1	9	0.344	0.969	0.988	35.69
2	"	0.522	0.834	0.902	36.56
3	"	0.625	0.744	0.838	36.70
4	"	0.741	0.623	0.755	46.50
5	"	0.778	0.581	0.713	44.88
6	"	0.803	0.552	0.719	42.83
7	"	0.825	0.524	0.748	41.72
8	"	0.842	0.502	0.786	40.62
<b>9</b>	<b>9</b>	<b>0.854</b>	<b>0.486</b>	<b>0.824</b>	<b>39.08</b>

SECV=0.713; F=9 SECV=0.824), which may be indicative of possible over-fitting [16]. The calibration was, however, a preliminary test to assess the samples collected for the final calibration from one season.

#### 4.6.4 Combined Seasons calibrations of CIE L\*

Table 4.14 Laboratory parameters summary of combined L\*

# Samples (in calibration)	# Samples in validation	Range	Mean	Standard Deviation
311(249)	62	27.23-37.17	32.54	1.75

##### 4.6.4.1 MLR Calibration

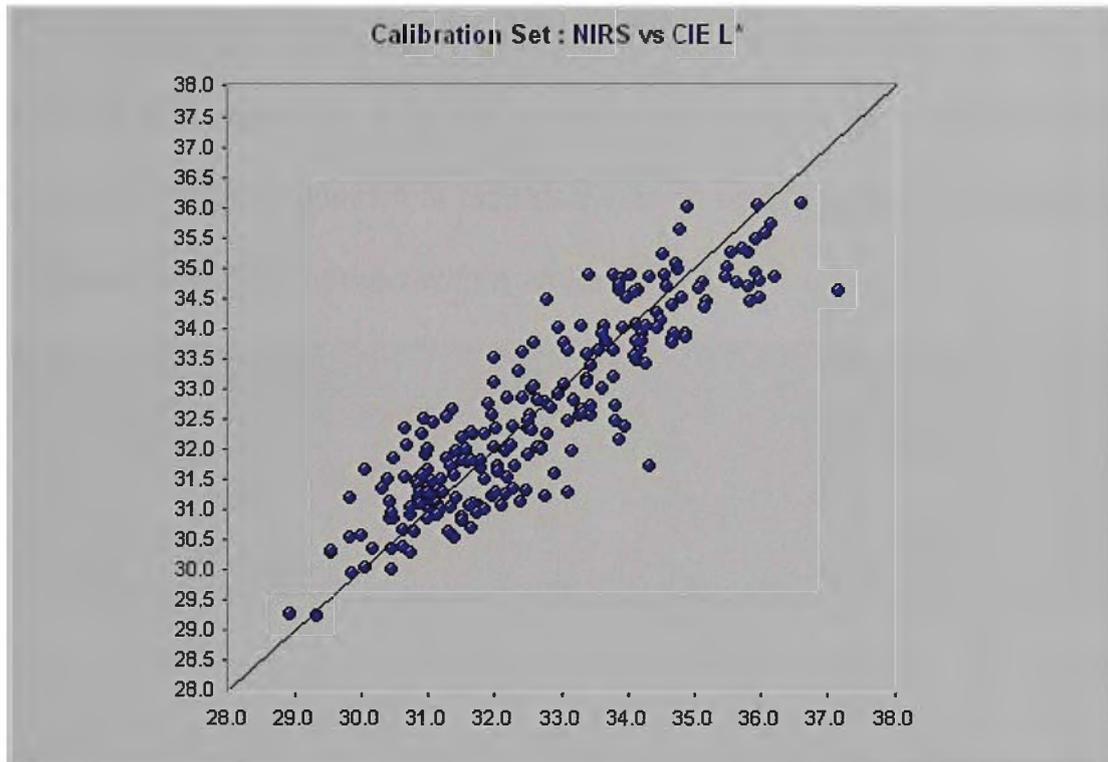


Figure 4.8 A plot of the combined seasons calibration for CIE L\* using MLR.

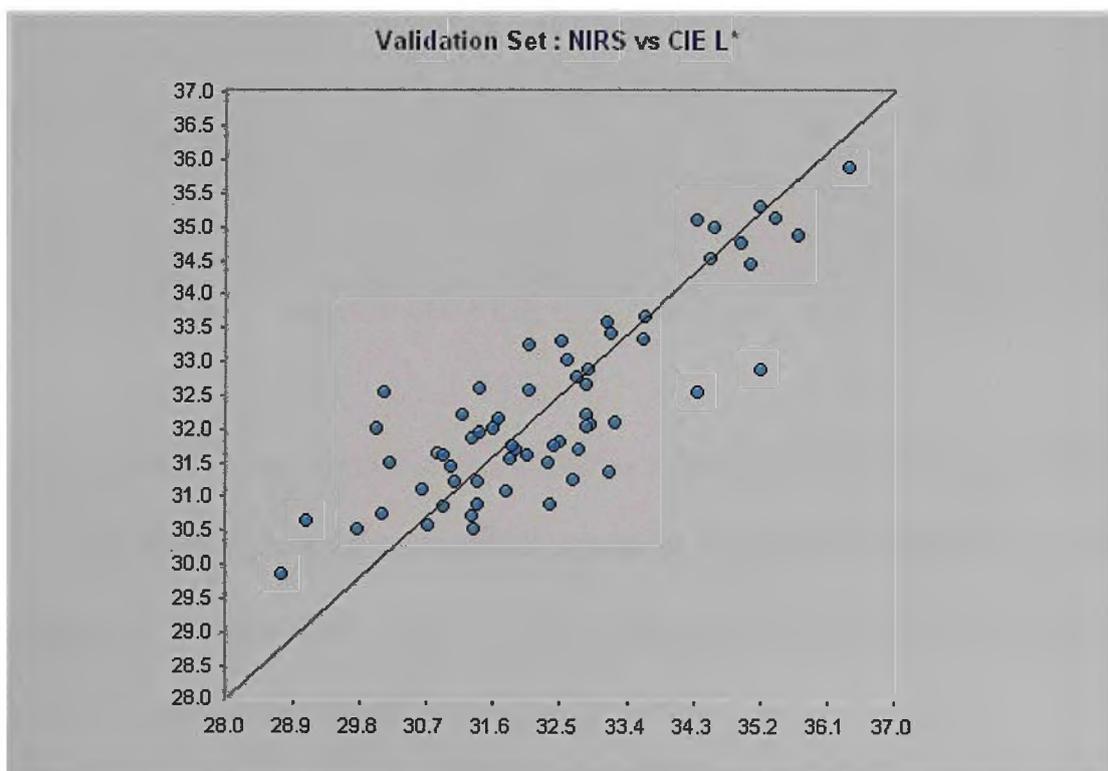


Figure 4.9 A plot of the combined seasons validation set for CIE L\* using MLR.

Included are diagrams illustrating three spectra from the calibration set (Figure 4.10), these are whole spectrum including the visible region as well as the NIR region (400 to 2500nm). The spectra correspond to the sample with the highest CIE L\* in red (L\*=37.17) the spectra of the sample closest to the mean CIE L\* in yellow (L\*=32.54) and the lowest sample in blue (L\*=27.23). A close up of the region of the first wavelength used in this MLR calibration (Figure 4.11) of those spectra is included with the point where the wavelength intercepts the spectra marked with a white line.

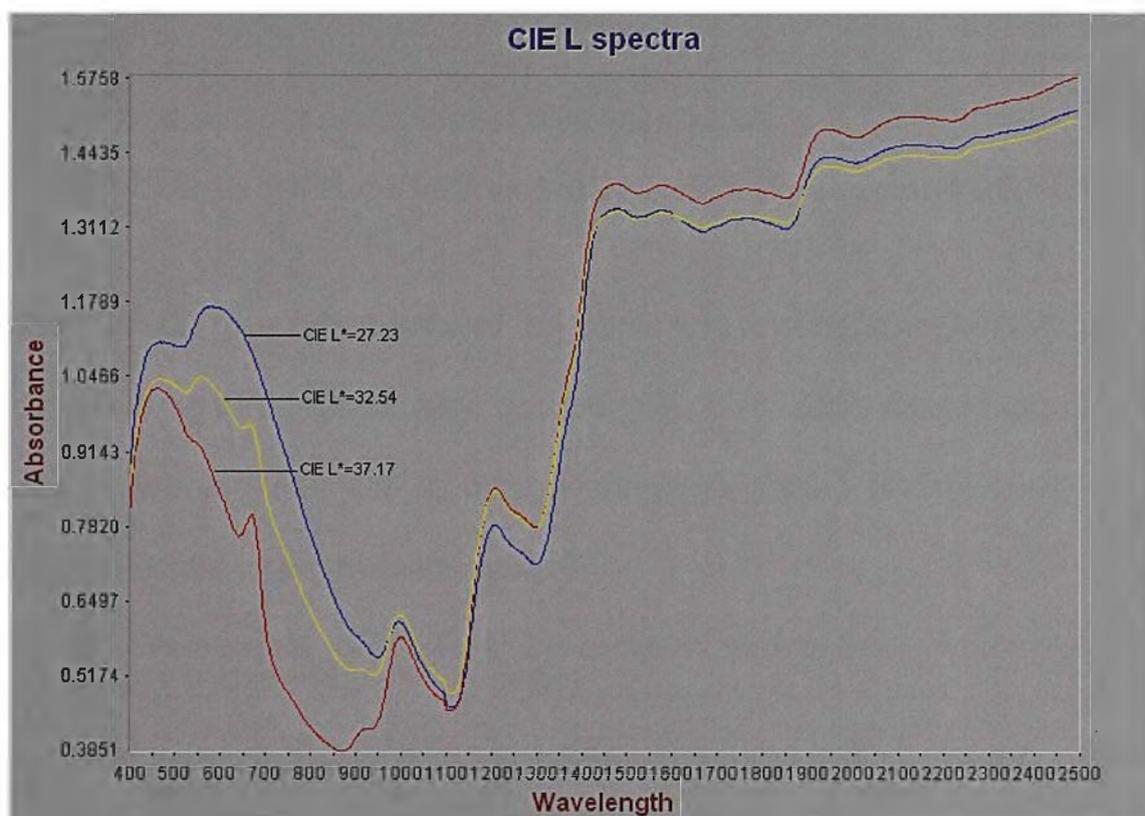


Figure 4.10 Spectrum of the minimum (blue), maximum (red) and mean (yellow) samples of CIE L\*

In contrast to some of the calibrations of the individual seasons, the combined CIE L\* (the light-dark axis of the CIE system) calibration displayed a strong association with the visible region of the spectrum (Table 4.15), four of the six summation terms originated from this region [12].

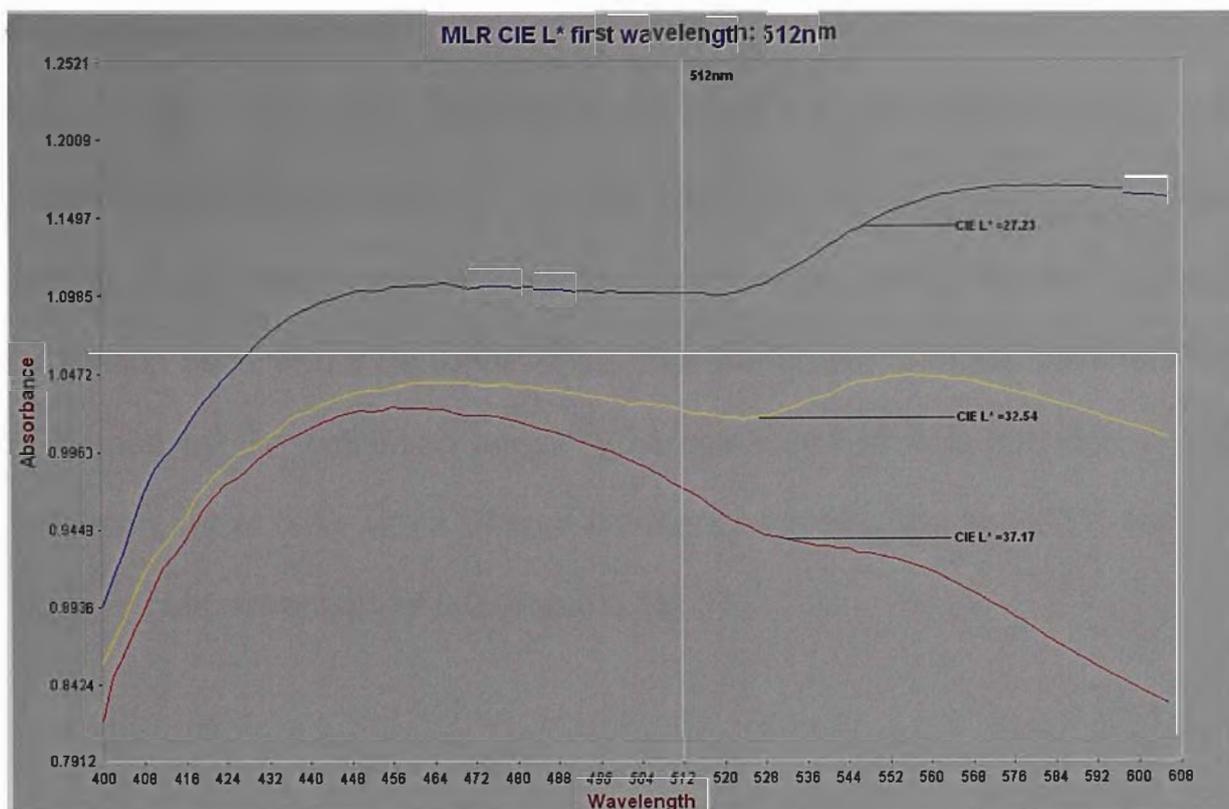


Figure 4.11 Close up of NIR spectra at the first wavelength of the combined MLR calibration.

Factor two (2318nm) had links to the stretching and deformation vibrations of carbon hydrogen bonds of alkyl groups [13]. The fourth wavelength (800nm) of this calibration was due to third overtone stretching vibrations of nitrogen hydrogen bonds of primary amines [13].

Table 4.15 Wavelengths contributing to the combined seasons MLR calibration of CIE L\* (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

MLR								
Term	Outliers	$\lambda$ (nm)	R <sup>2</sup> (Cal)	SEC	F-value	R <sup>2</sup> (Val)	Bias	SEP
1	10	512	0.619	1.06	357.4	-	-	-
2	"	2318	0.706	0.935	262.5	-	-	-
3	"	530	0.736	0.888	202.5	-	-	-
4	"	800	0.767	0.837	178.1	-	-	-
5	"	566	0.788	0.799	160.4	-	-	-
<b>6</b>	<b>10</b>	<b>456</b>	<b>0.799</b>	<b>0.780</b>	<b>142.5</b>	<b>0.840</b>	<b>0.0465</b>	<b>0.900</b>

Included above are plots of the calibration set (Figure 4.8) and the validation set (Figure 4.9) of this calibration. As more terms were added to the calibration the coefficient of determination as well as SEC showed clear

improvement (Table 4.15 Term=1  $R^2= 0.619$ , SEC= 1.06; Term=6  $R^2= 0.799$ , SEC= 0.78). The overall trend seen of the F-value for this calibration was one of steady decrease (Table 4.15 F-value trend: 357 to 143), however the final size of the F-value was comparatively large which would indicate that the calibration itself would be robust [15]. The performance of the calibration is confirmed by the validation achieving a coefficient of determination of the validation set of 0.84 with a Bias of 0.0465 and an uncorrected SEP, without slope or bias correction, of 0.9 (Table 4.15).

#### 4.6.4.2 PLS Calibration

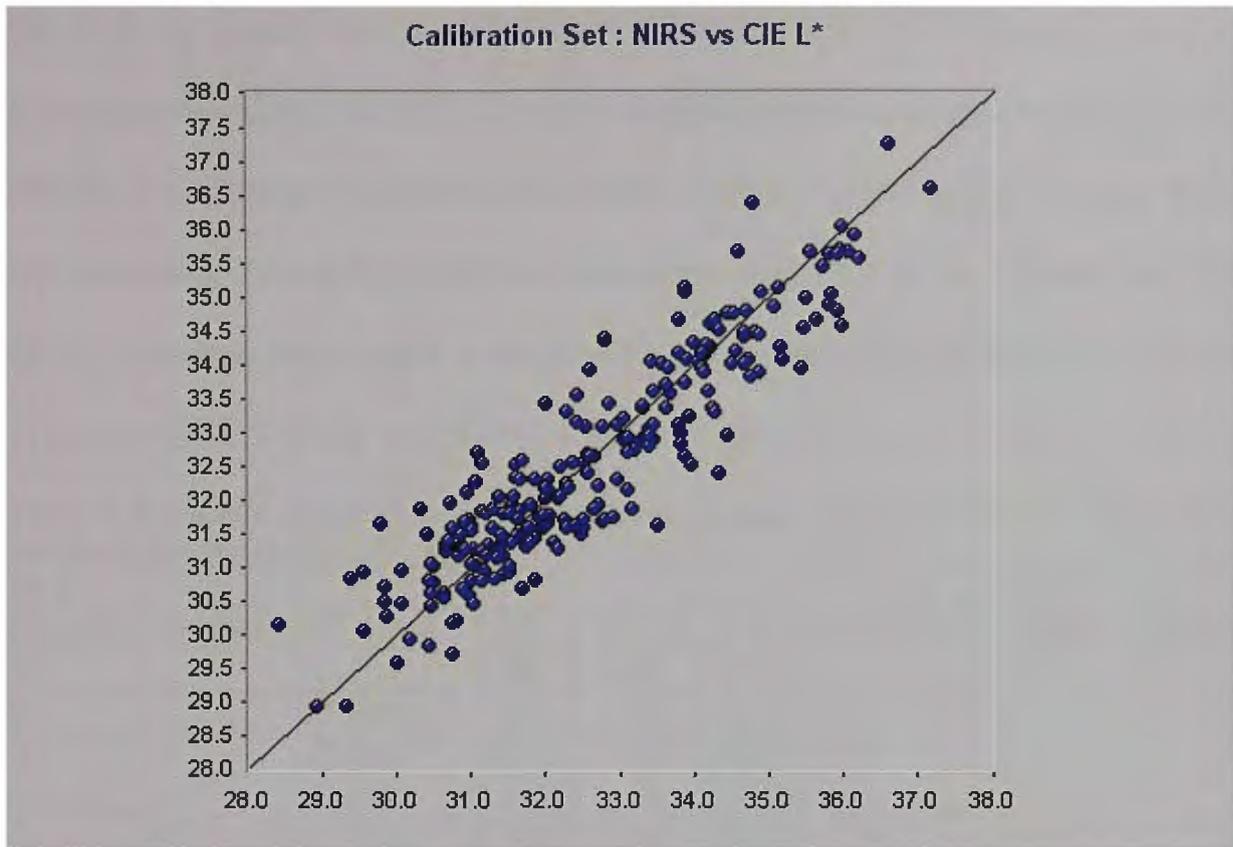


Figure 4.12 A plot of the combined seasons calibration for CIE L\* using PLS.

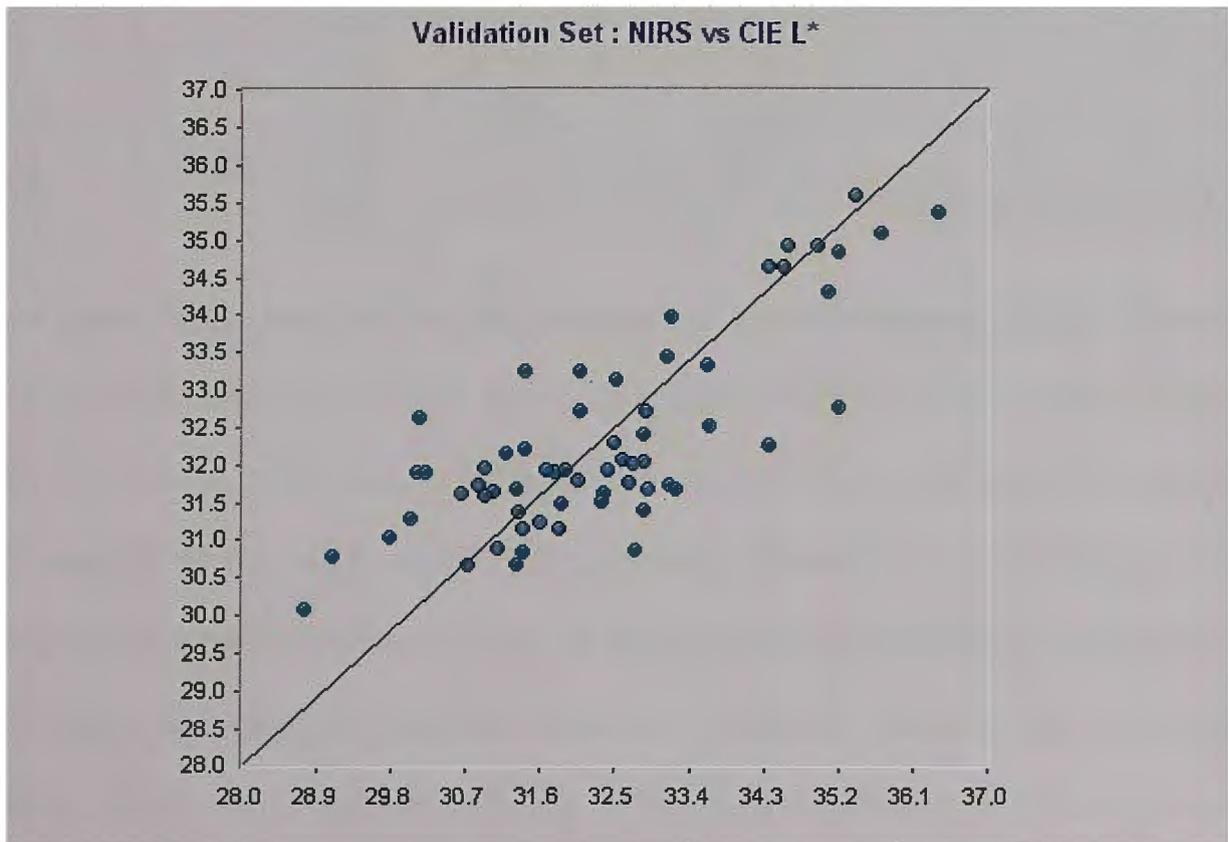


Figure 4.13 A plot of the combined seasons validation set for CIE L\* using PLS

Plots of the calibration set (Figure 4.12) and the validation set (Figure 4.13) of the PLS calibration are provided above. Both PLS and MLR were used to develop calibrations for CIE L\*, each method was successful in this model, with PLS achieving the best results, particularly considering that outliers were not removed, indicating that this calibration would be more robust than the MLR calibration which gave a correlation only after the removal of ten outliers (Table 4.16 PLS  $R^2= 0.835$ ; Table 4.15 MLR  $R^2= 0.799$ ).

Table 4.16 Factors contributing to the combined seasons PLS calibration of CIE L\* (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

PLS								
Factors	Outliers	$R^2$ (Cal)	SEC	SECV	F-value	$R^2$ (Val)	Bias	SEP
1	-	0.484	1.26	1.26	215.5	-	-	-
2	-	0.644	1.04	1.12	207.6	-	-	-
3	-	0.680	0.993	1.04	161.8	-	-	-
4	-	0.696	0.970	1.04	129.8	-	-	-
5	-	0.708	0.953	1.04	109.5	-	-	-
6	-	0.750	0.883	1.01	112.7	-	-	-
7	-	0.766	0.857	0.992	104.7	-	-	-
8	-	0.790	0.813	0.956	105.2	-	-	-
9	-	0.806	0.783	0.945	102.6	-	-	-
10	-	0.816	0.765	0.941	97.92	-	-	-
11	-	0.825	0.747	0.941	94.43	-	-	-
<b>12</b>	-	<b>0.835</b>	<b>0.727</b>	<b>0.940</b>	<b>92.61</b>	<b>0.792</b>	<b>0.069</b>	<b>1.05</b>

As more factors were added the coefficient of determination and SEC showed clear improvement (Factor=1  $R^2= 0.484$ , SEC= 1.26; Factor= 12  $R^2= 0.835$ , SEC= 0.727). The F-value (Table 4.16 F-value of PLS= 92.6 compared to an F-value of MLR= 142.5) of the PLS calibration showed a trend contrary to this argument, when compared PLS to MLR, however, if this value was compared to those obtained in individual seasons, a marked improvement could be seen, which was a positive indicator of improved reliability [15]. The standard error of cross validation values showed a clear trend of gradual decline towards a minimum which was reached at the final factor of the calibration

(Table 4.16  $F= 12$   $SECV= 0.94$ ), gave a clear indication that the calibration was robust [17]. The coefficient of determination for the validation set achieved was 0.792 with a Bias of 0.069 and an uncorrected SEP of 1.05 (Table 4.16).

## 4.7 Tricolour Stimulus a\*

### 4.7.1 Season one CIE a\*

Table 4.17 Laboratory parameters summary of season one a\*

# Samples (in calibration)	# Samples in validation	Range	Mean	Standard Deviation
111 (89)	22	0.309-6.36	5.41	0.488

#### 4.7.1.1 MLR calibration

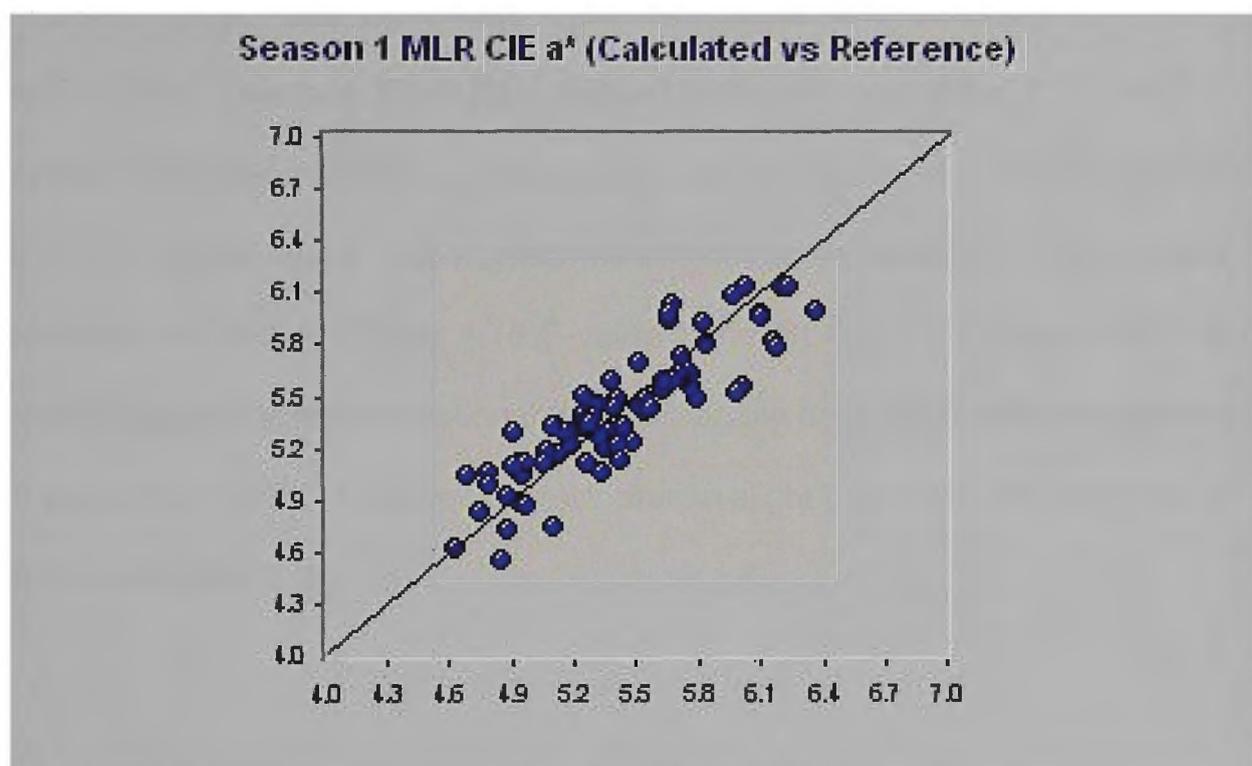


Figure 4.14 A plot of the first seasons calibration for CIE a\* using MLR

The MLR terms used to develop the calibration of CIE a\* for the first season were mostly attributable to the visible region of the NIRS spectrum, except for the third term (Table 4.18) [12]. The third term wavelength (1830nm) was linked to oxygen hydrogen bond and carbon-oxygen bond stretching vibrations [13], typically attributed to cellulose [14] as well as alkyl and terminal alkenyl groups [13]. Provided above is a plot of the MLR calibration set (Figure 4.14).

Table 4.18 Wavelengths contributing to the first seasons MLR calibration of CIE a\* (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

MLR					
Term	Outliers	$\lambda$ (nm)	R <sup>2</sup> (Cal)	SEC	F-value
1	9	597	0.455	0.309	61.82
2	"	462	0.601	0.267	54.97
3	"	1830	0.675	0.242	49.77
4	9	488	0.738	0.219	49.89

As more terms were added to the MLR a\* calibration set, gains in coefficient of determination and SEC were achieved (Table 4.18 Terms=1 R<sup>2</sup>=0.433, SEC=0.309; Terms=4 R<sup>2</sup>=0.739, SEC=0.219). It was difficult to build a suitable calibration model of a\* due to the narrow range of samples and the fact that greenness is not a predominant colour in sultanas. The steady decrease in F-value (Table 4.18 F-value trend=61.8 to 49.9) seen over the calibration gave some indication that the reliability decreased with the addition of summation terms, however the rate was so slight that it was not detrimental to the calibration [15].

#### 4.7.1.2 PLS calibration

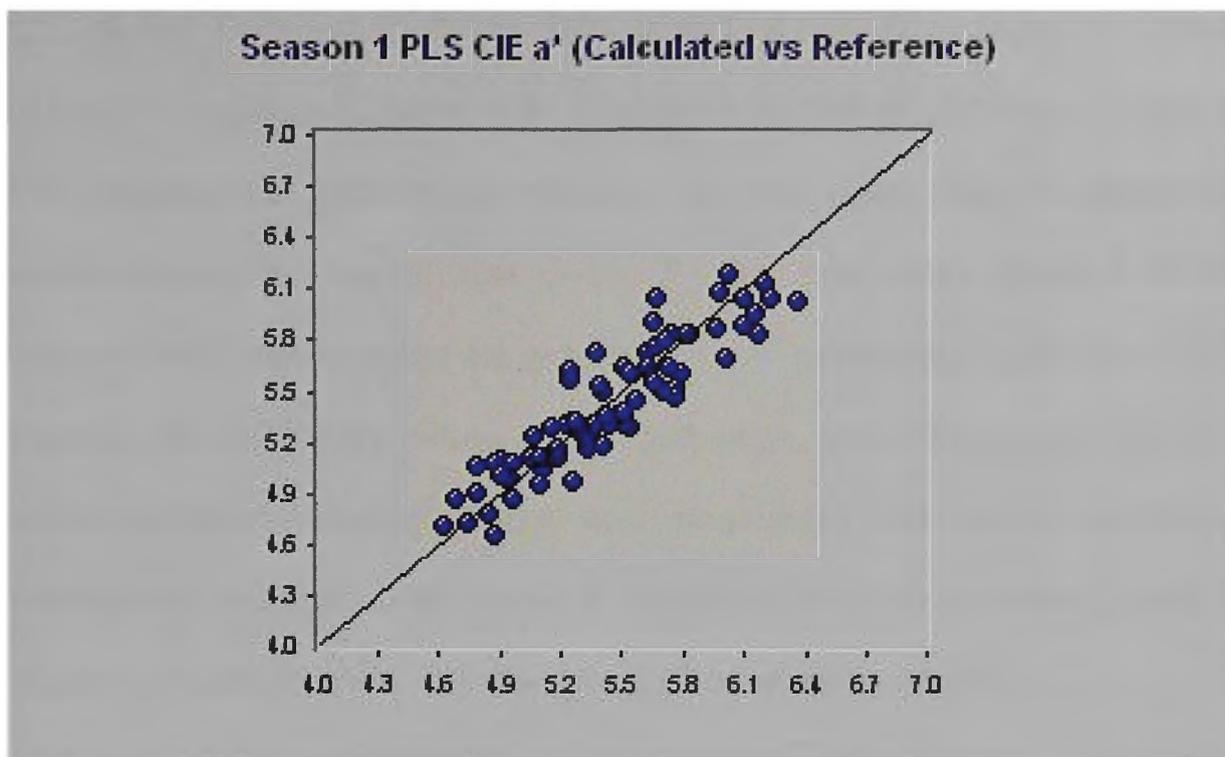


Figure 4.15 A plot of the first seasons calibration for CIE a\* using PLS

In contrast with the corresponding MLR spectrum, PLS resulted in a more robust calibration, bearing in mind the sample number and seasonal variation of the calibration. Above is provided a plot of the PLS calibration (Figure 4.15).

Table 4.19 Factors contributing to the first seasons PLS calibration of CIE a\* (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

PLS					
Factors	Outliers	R <sup>2</sup> (Cal)	SEC	SECV	F-value
1	9	0.305	0.349	0.358	32.55
2	"	0.415	0.323	0.354	25.89
3	"	0.537	0.289	0.355	27.83
4	"	0.636	0.258	0.341	31.00
5	"	0.696	0.237	0.330	32.11
6	"	0.729	0.226	0.313	30.91
7	"	0.776	0.207	0.303	33.59
8	"	0.798	0.198	0.300	33.14
<b>9</b>	<b>9</b>	<b>0.836</b>	<b>0.180</b>	<b>0.300</b>	<b>37.46</b>

As more factors were introduced into the calibration to explain variance between the spectroscopic and laboratory data, the coefficient of

determination and SEC improved (Table 4.19 Factor=1  $R^2=0.305$ , SEC=0.349; Factor=9  $R^2=0.839$ , SEC=0.180). A general increase in F-value of the PLS calibration (Table 4.19 F-value trend=32.6 to 37.5) was observed, this showed the calibrations reliability did not suffer due to calibration development. The overall size of the F-value was small (Table 4.19 F-value=37.46), which would be expected for a preliminary calibration [15]. Despite the preliminary nature of this calibration, the trend seen in error of cross validation indicated that it was progressing well, as it reached a minimum at the final factor (Table 4.19 F=9 SECV=0.30) indicating that the addition of further factors not affect the calibration adversely [17].

## 4.7.2 Season two CIE a\*

Table 4.20 Laboratory parameters summary of season two a\*

# Samples (in calibration)	# Samples in validation	Range	Mean	Standard Deviation
95 (76)	19	3.41-6.64	5.04	0.633

### 4.7.2.1 MLR Calibration

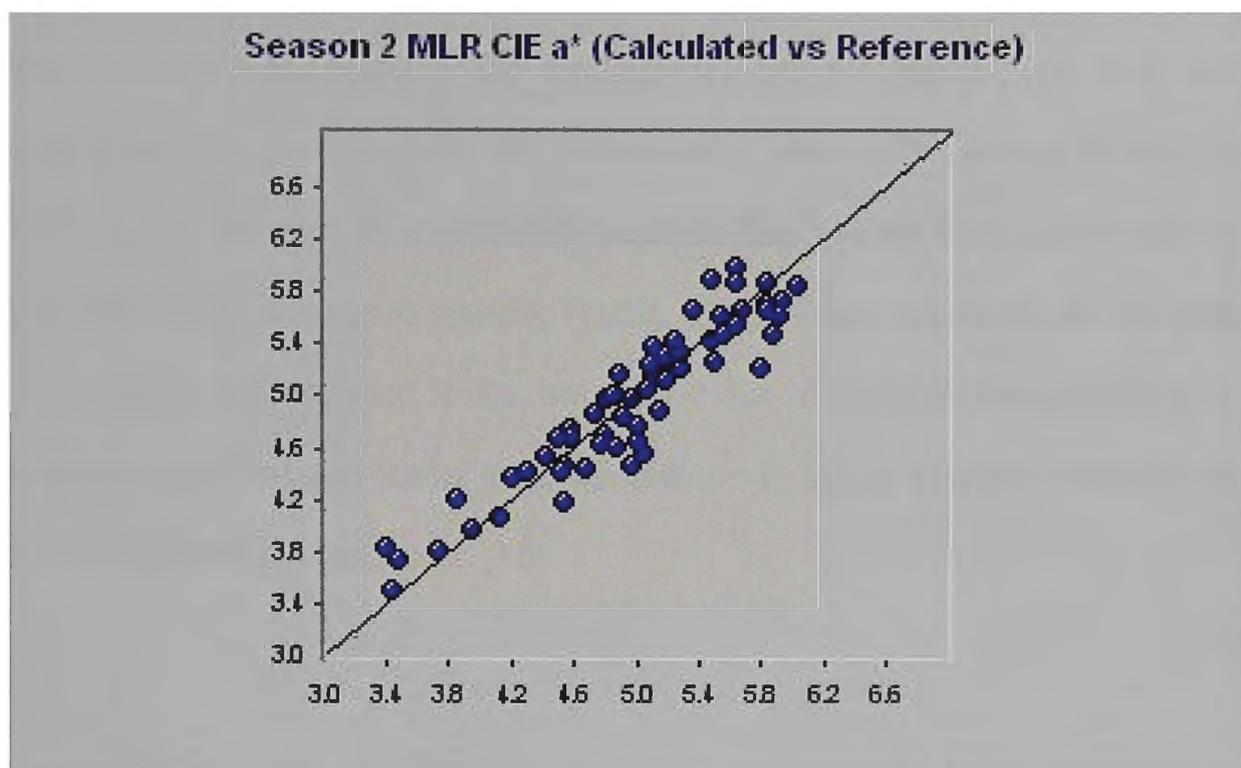


Figure 4.16 A plot of the second seasons calibration for CIE a\* using MLR

The first two terms (470nm and 490nm) used to develop a calibration for CIE a\* were in the visible region of the spectra [13]. The third term (1972nm) was attributed to a trend seen in a region of the spectra due to the asymmetric stretching vibrations of nitrogen-hydrogen bonds and in secondary amides [12], often attributed to protein [13] as well as the carbonyl of ketones, and water, carboxylic acids, phenols and tertiary hydroxy groups [14]. The fourth term (784nm) is from a region that can be considered to be from stretching vibrations of nitrogen-hydrogen bonds [12] often seen when primary amines are present [14]. A plot of the MLR calibration set is provided above (Figure 4.16).

Table 4.21 Wavelengths used to contribute to the second seasons MLR calibration for CIE a\* (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

MLR					
Term	Outliers	$\lambda$ (nm)	R <sup>2</sup> (Cal)	SEC	F-value
1	8	470	0.672	0.372	120.65
2	"	494	0.733	0.338	79.71
3	"	1972	0.772	0.316	64.3
<b>4</b>	<b>8</b>	<b>784</b>	<b>0.839</b>	<b>0.268</b>	<b>72.98</b>

As more summation terms were added to the MLR calibration, steady improvement was seen in the performance of the calibration, both the coefficient of correlation and SEC improved (Table 4.21 Term=1 R<sup>2</sup>=0.672, SEC=0.372; Term=4 R<sup>2</sup>=0.839, SEC=0.268) The overall trend in F-value of this calibration decreased steadily (Table 4.21 F-value trend=121 to 73) with the addition of summation terms, however it was of sufficient size (Table 4.21 F-value=73) in relative terms to be considered to give a positive indication of the reliability of the calibration [15].

#### 4.7.2.2 PLS Calibration

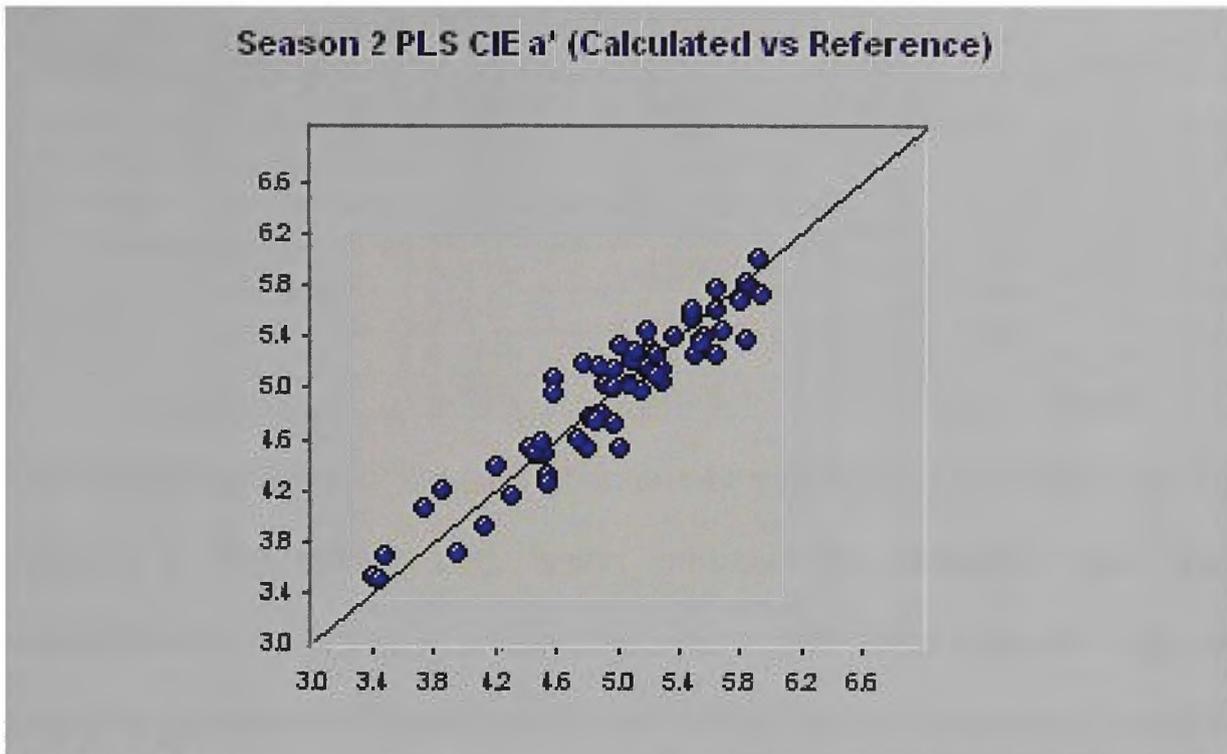


Figure 4.17 A plot of the second seasons calibration for CIE a\* using PLS

A plot of the PLS calibration set is provided above (Figure 4.17). When contrasting the PLS and MLR calibrations of the second season, the performance of PLS over MLR shows a clear improvement in both its correlation and error of calibration (Table 4.21 MLR  $R^2=0.839$ ,  $SEC=0.268$ ; Table 4.22 PLS  $R^2=0.885$ ,  $SEC=0.226$ ). This improvement in performance was compensated by a slight reduction in reliability as shown by the F-value, (Table 4.21 MLR F-value=73; Table 4.22 PLS F-value=51.1) however, this was a less marked reduction than in previous examples [15].

As more factors were included into the calibration to explain variance between spectroscopic data and laboratory data, the coefficient of determination and SEC improved (Table 4.22 Factor=1  $R^2=0.574$ ,  $SEC=0.410$ ; Factor=8  $R^2=0.885$ ,  $SEC=0.226$ ). The trend in the F-values (Table 4.22 F-value trend=80.76 to 51.1) of this calibration showed a gradual decrease, however the rate of decrease was not sufficient to affect this calibration [15].

Table 4.22 Factors contributing to second seasons PLS calibration of CIE a\* (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

PLS					
Factors	Outliers	R <sup>2</sup> (Cal)	SEC	SECV	F-value
1	8	0.574	0.410	0.419	80.76
2	"	0.629	0.386	0.416	50.02
3	"	0.721	0.337	0.407	50.08
4	"	0.754	0.319	0.404	43.79
5	"	0.814	0.280	0.388	49.05
6	"	0.844	0.259	0.393	49.49
7	"	0.859	0.249	0.386	46.91
<b>8</b>	<b>8</b>	<b>0.885</b>	<b>0.226</b>	<b>0.415</b>	<b>51.05</b>

The training set used for the calibration was of reasonable size (Table 4.20 76 samples in the training set), which indicated the calibration was less susceptible to over-fitting. The standard error of calibration indicated that this preliminary calibration risked over-fitting if further factors were added, as little improvement was seen, only some local minima was seen in the in SECV values at the fifth (Table 4.22 SECV=0.388) and seventh factors (Table 4.22 SECV=0.386) of the calibration [17].

### 4.7.3 Season three CIE a\*

Table 4.23 Laboratory parameters summary of season three CIE a\*

# Samples (in calibration)	# Samples in validation	Range	Mean	Standard Deviation
108 (86)	22	3.06-6.77	5.28	0.64

#### 4.7.3.1 MLR Calibration

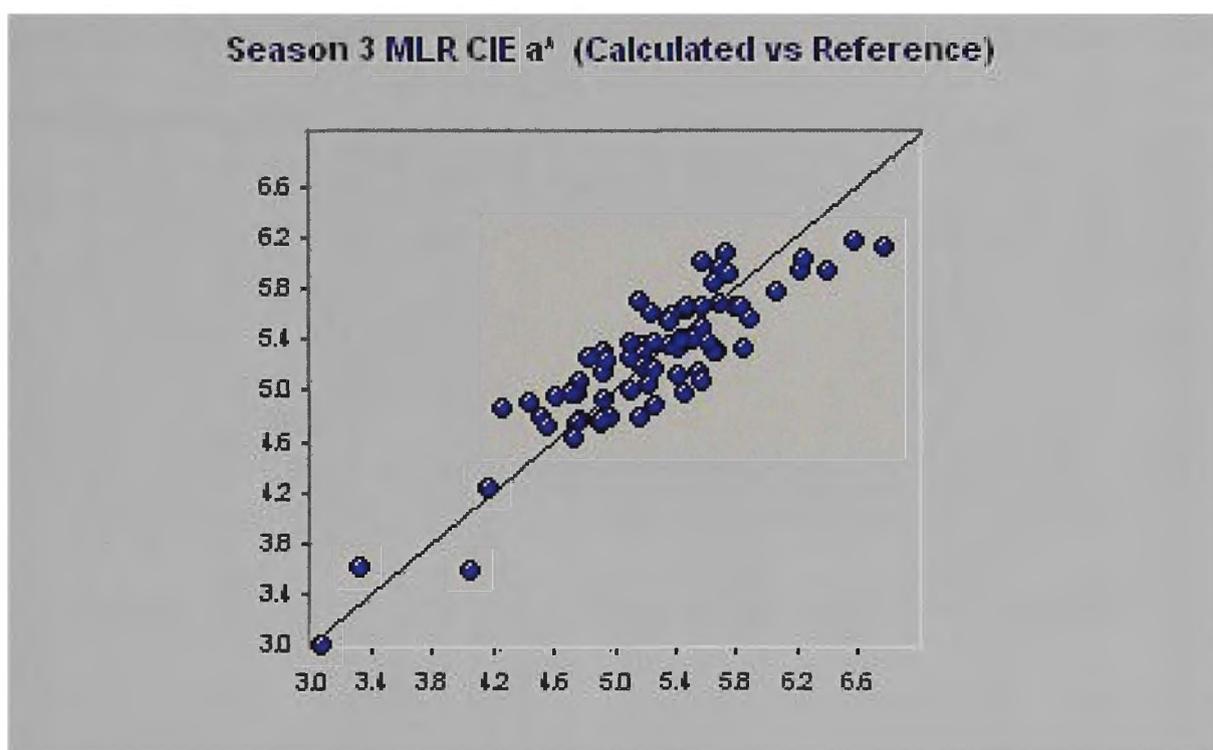


Figure 4.18 A plot of the third seasons calibration for CIE a\* using MLR

In contrast to the calibration of L\*, a\* showed a significant link to the visible region of the spectra; only the second term originated outside the visible region of the spectrum [12]. The second term (Table 4.24 2318nm) can be attributed to carbon hydrogen bond stretching and deformation vibrations of alkyl and aryl groups [14] from the combinations band [12]. As more terms were included into the calibration to explain variance between the laboratory reference data and spectroscopic data, the coefficient of determination and SEC steadily improved (Table 4.24 Term=1  $R^2=0.417$ , SEC=0.494; Term=4  $R^2=0.793$ , SEC= 0.301). The general trend in the F-value (Table 4.24 F-

Table 4.24 Wavelengths contributing to the third seasons calibration MLR of CIE a\* (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

MLR					
Term	Outliers	$\lambda$ (nm)	R <sup>2</sup> (Cal)	SEC	F-value
1	9	590	0.417	0.494	48.59
2	"	2318	0.621	0.401	54.88
3	"	544	0.766	0.318	72.03
4	9	516	0.793	0.301	62.23

value=48.6 to 62.2) of this calibration showed a steady increase, it was also large enough to (Table 4.24 F-value=62.23) indicate the calibration would predict reliably [15].

### 4.7.3.2 PLS Calibration

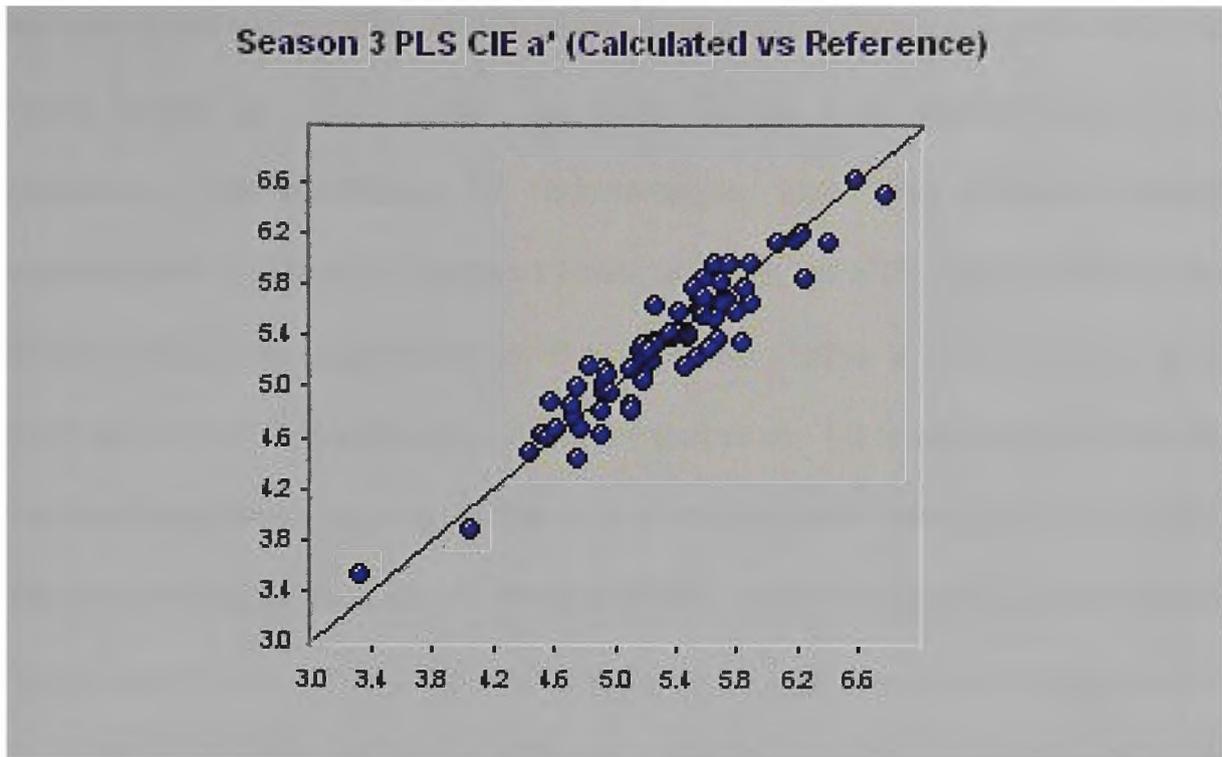


Figure 4.19 A plot of the third seasons calibration for CIE a\* using PLS

Provided above is a plot of the PLS calibration (Figure 4.19). The observed performance of both calibrations of the third season, showed a clear improvement in performance of PLS over MLR. The error of calibration and coefficient of determination were significantly improved (Table 4.25 PLS  $R^2=0.89$ ,  $SEC=0.20$ ; 4.24 MLR  $R^2=0.793$ ,  $SEC=0.30$ ), while the difference in the F-value was minimal (Table 4.24 MLR F-value=62.23; Table 4.25 PLS F-value=57.16).

Table 4.25 Factors contributing to the third seasons PLS calibration for CIE a\* (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

PLS					
Factors	Outliers	$R^2$ (Cal)	SEC	SECV	F-value
1	9	0.1461	0.533	0.550	11.81
2	"	0.407	0.448	0.519	23.35
3	"	0.527	0.403	0.451	24.89
4	"	0.592	0.377	0.443	23.91
5	"	0.737	0.305	0.381	36.48
6	"	0.807	0.263	0.343	44.69
7	"	0.857	0.228	0.316	53.90
8	"	0.880	0.210	0.300	57.09
<b>9</b>	<b>9</b>	<b>0.894</b>	<b>0.200</b>	<b>0.294</b>	<b>57.16</b>

Considering the fact that this calibration was preliminary as it contained samples from one season, so when the combined calibration is optimised, the model would be very reliable. As more factors were added to the PLS calibration, the coefficient of determination and SEC showed steady improvement (Table 4.25 Factor=1  $R^2=0.146$ , SEC=0.533; Factor=9  $R^2=0.894$  SEC=0.200). The trend seen for the F-values (Table 4.25 F-value=11.8 to 57.2) within the PLS calibration mirrored that of the MLR calibration however; the trend was more marked, as the rate of improvement was greater. This and the relatively large size of the F-value (Table 4.25 F-value=57.2) indicated a reliable calibration [15]. Clear improvement in the error of cross validations as each factor was added to the calibration was seen without reaching a clear minimum before the final factor was reached (Table 4.25 Factor=1 SECV=0.55; Factor=9 SECV=0.294), which was satisfactory for a preliminary calibration [17].

#### 4.7.4 Combined seasons CIE a\*

Table 4.26 Laboratory parameters summary of combined a\*

# Samples (in calibration)	# Samples in validation	Range	Mean	Standard Deviation
311(249)	62	3.06-6.77	5.26	0.607

##### 4.7.4.1 MLR Calibration

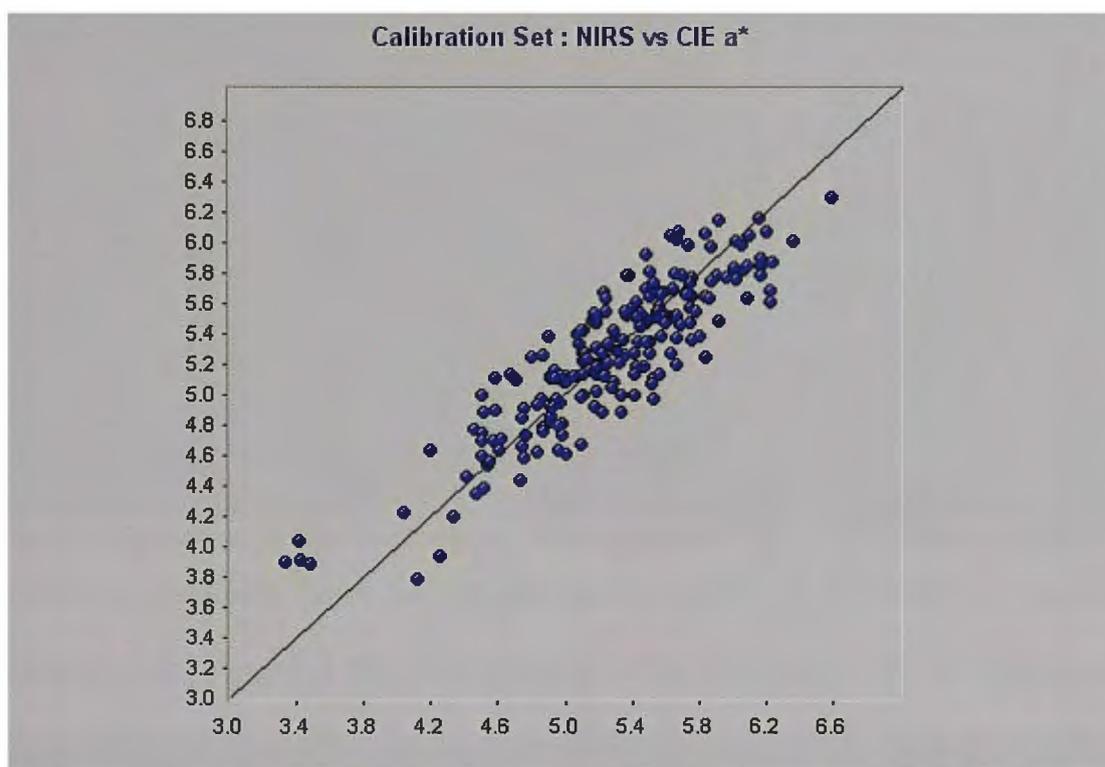


Figure 4.20 A plot of the combined seasons calibration for CIE a\* using MLR

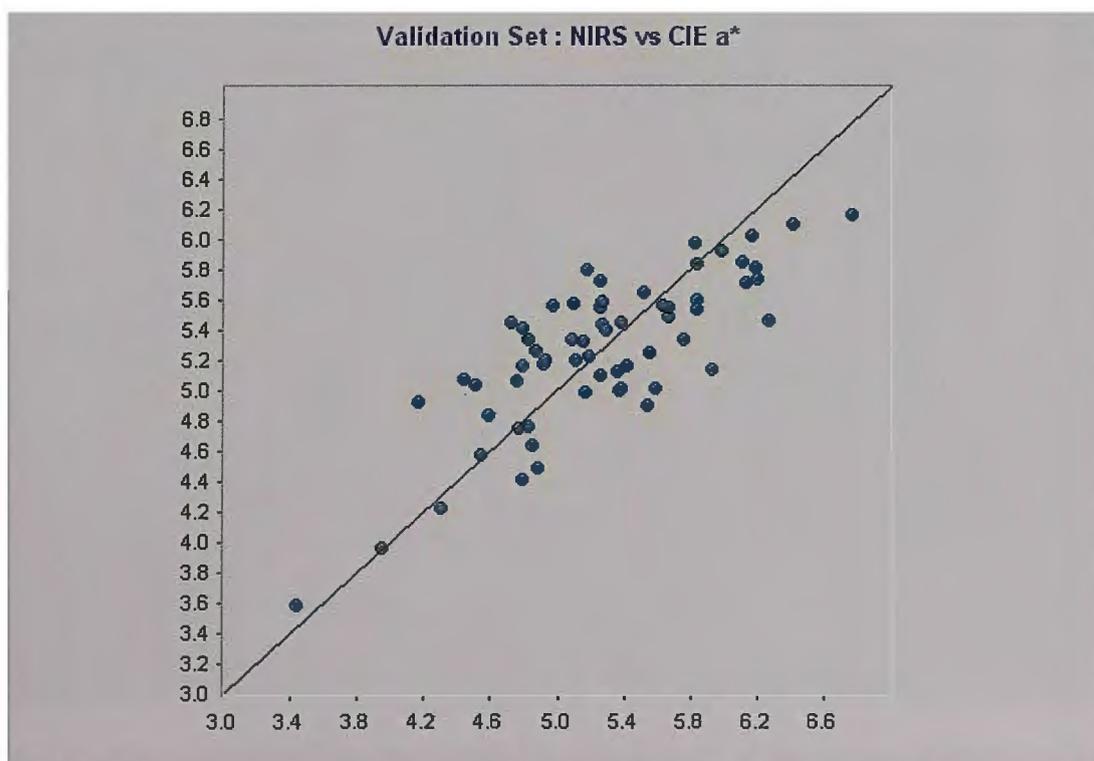


Figure 4.21 A plot of the combined seasons validation set for CIE a\* using MLS

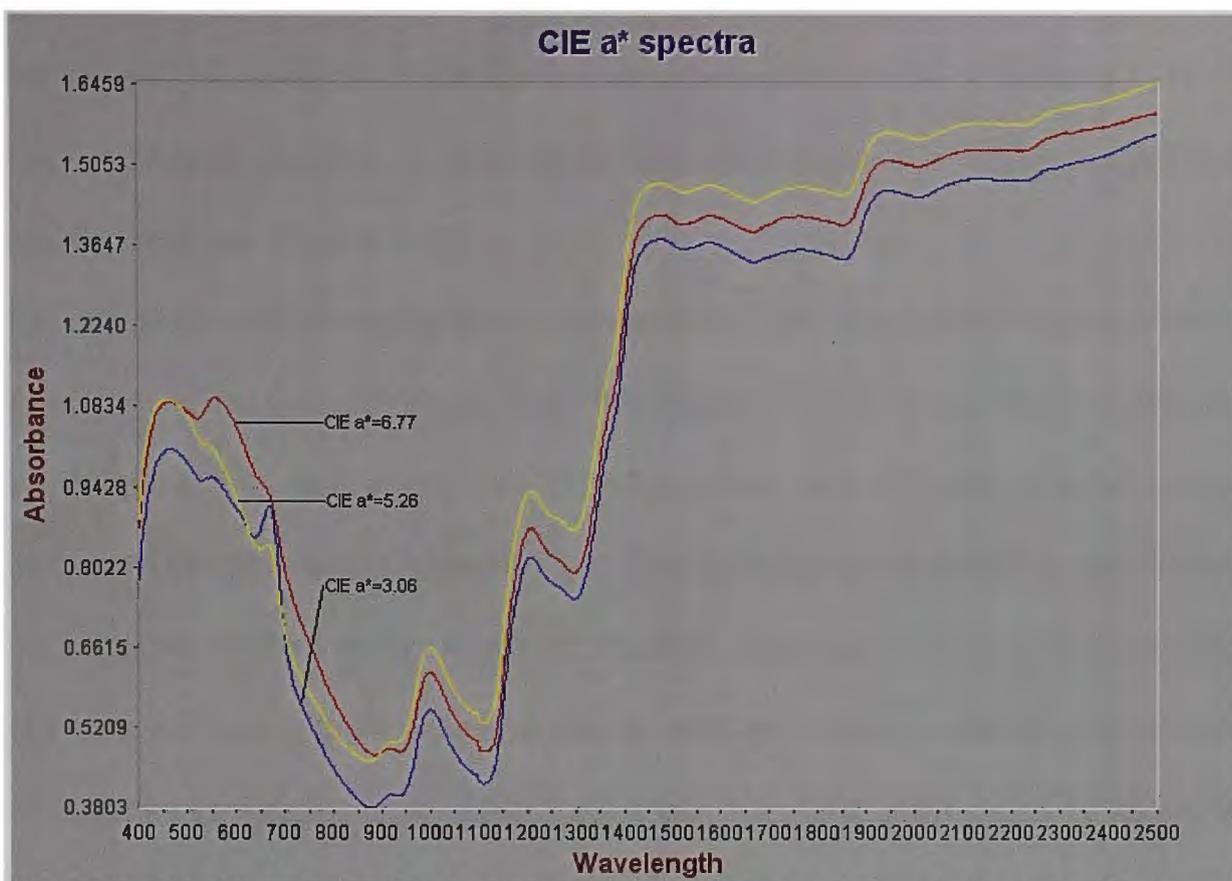


Figure 4.22 Spectra of the minimum (blue), maximum (red) and mean (yellow) of CIE a\* Provided above are three full range spectra (400 to 2500nm) of the CIE a\* calibration set (Figure 4.22), the spectra of the maximum CIE a\* value sample

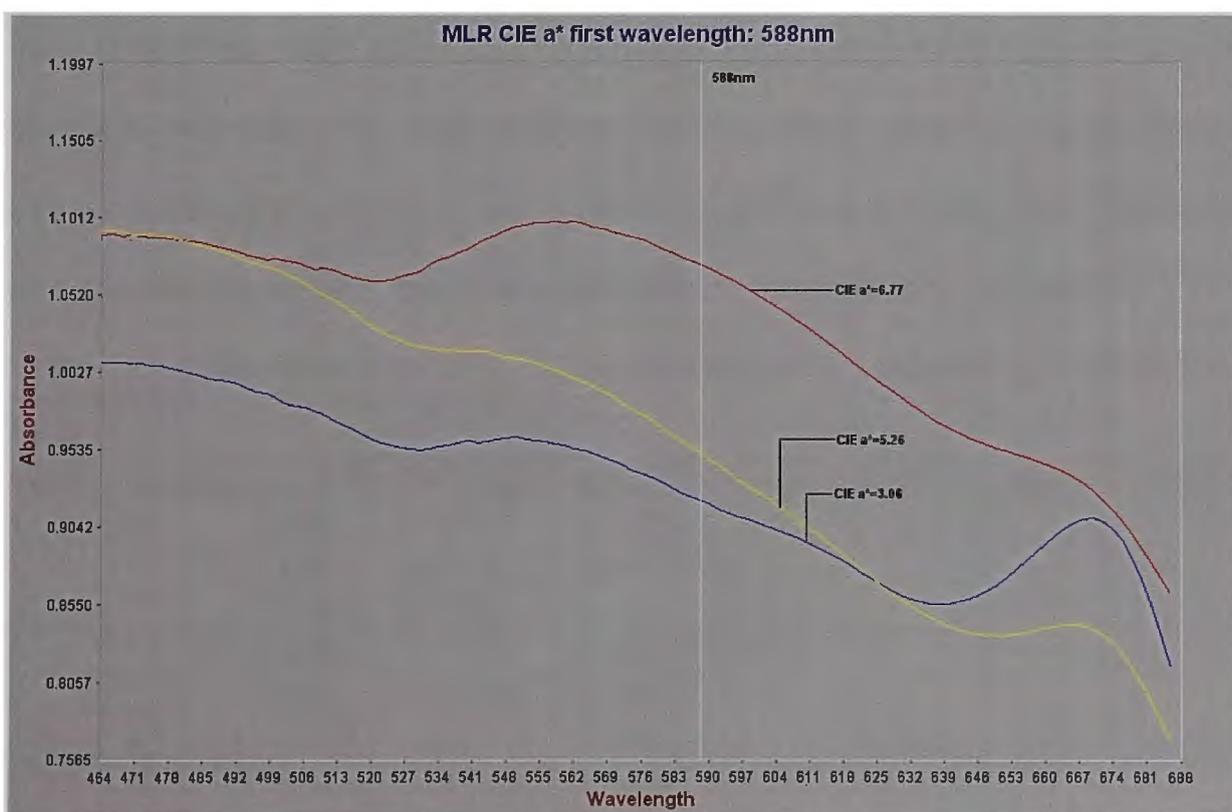


Figure 4.23 Close up of NIR spectra at the first wavelength of the combined MLR calibration

(6.77) in red, the spectra of the sample closest to the mean value (5.26) in yellow and the spectra of the sample with the minimum CIE a\* value (3.06) in blue. Included above is a close up of the region of the first wavelength of the MLR calibration (Figure 4.23).

CIE a\* also had a strong direct association with the visible region of the spectrum, three terms originate from this region [13]. The third term (1260nm) is attributable to first overtone [12] alkyl group and terminal alkenyl group carbon hydrogen bond vibrations [14]. The fourth term (1982nm) is due to the asymmetric stretching vibrations of nitrogen hydrogen bonds in protein [13] such as amides and primary amines, as well as terminal alkenyl groups and all hydroxy species except primary alcohols [14]. The sixth term (1832nm) is associated with the stretching vibrations of oxygen hydrogen and carbon oxygen bonds in cellulose [13], as well as carbon hydrogen groups of alkyl groups and alkenes with a hydrogen bonded to it [14]. Provided above are plots of the calibration set (Figure 4.20) and the validation set (Figure 4.21) of the MLR calibration. As more summation terms were added to the combined seasons calibration of CIE a\* to account for variance between the laboratory and spectroscopic data, both the coefficient of determination and SEC

Table 4.27 Wavelengths contributing to the combined MLR calibration for CIE b\* (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

MLR								
Term	Outliers	$\lambda$ (nm)	R <sup>2</sup> (Cal)	SEC	F-value	R <sup>2</sup> (Val)	Bias	SEP
1	25	588	0.364	0.437	113.4	-	-	-
2	"	484	0.576	0.360	130.8	-	-	-
3	"	1260	0.666	0.379	130.0	-	-	-
4	"	1982	0.710	0.298	119.2	-	-	-
5	"	488	0.7443	0.280	112.9	-	-	-
<b>6</b>	<b>25</b>	<b>1832</b>	<b>0.777</b>	<b>0.262</b>	<b>112.2</b>	<b>0.794</b>	<b>0.0657</b>	<b>0.386</b>

showed steady improvement (Table 4.27 Term=1 R<sup>2</sup>=0.364, SEC=0.437; Term=6 R<sup>2</sup>=0.777, SEC=0.262). The F-value during the development of this

calibration showed a general trend of decreasing size (Table 4.27 F-value=113 to 112), the rate of decrease was very slight and therefore indicated that the reliability of the calibration had not been adversely affected during development [15]. The coefficient of determination of the validation set achieved was 0.794 with an uncorrected SEP of 0.386 and a Bias of 0.0657.

#### 4.7.4.2 PLS Calibration

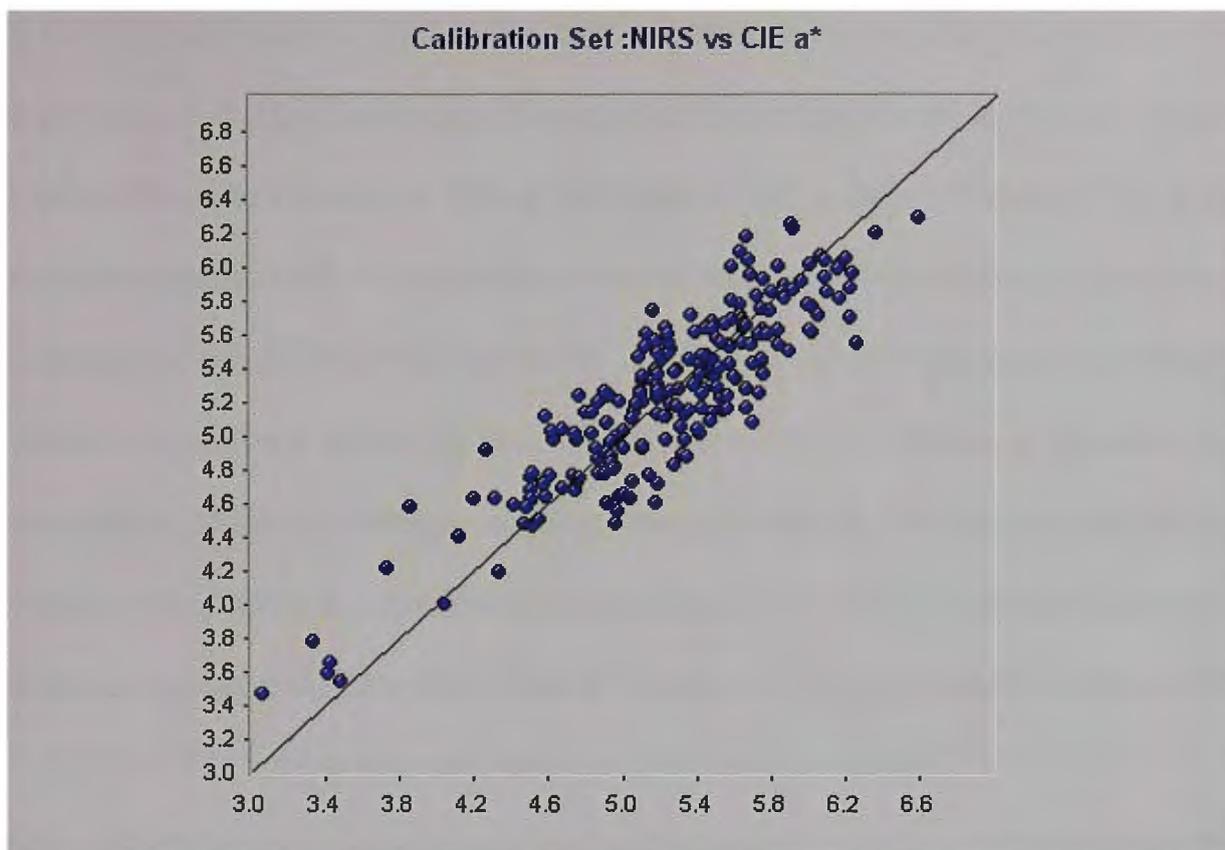


Figure 4.24 A plot of the combined seasons calibration for CIE a\* using PLS

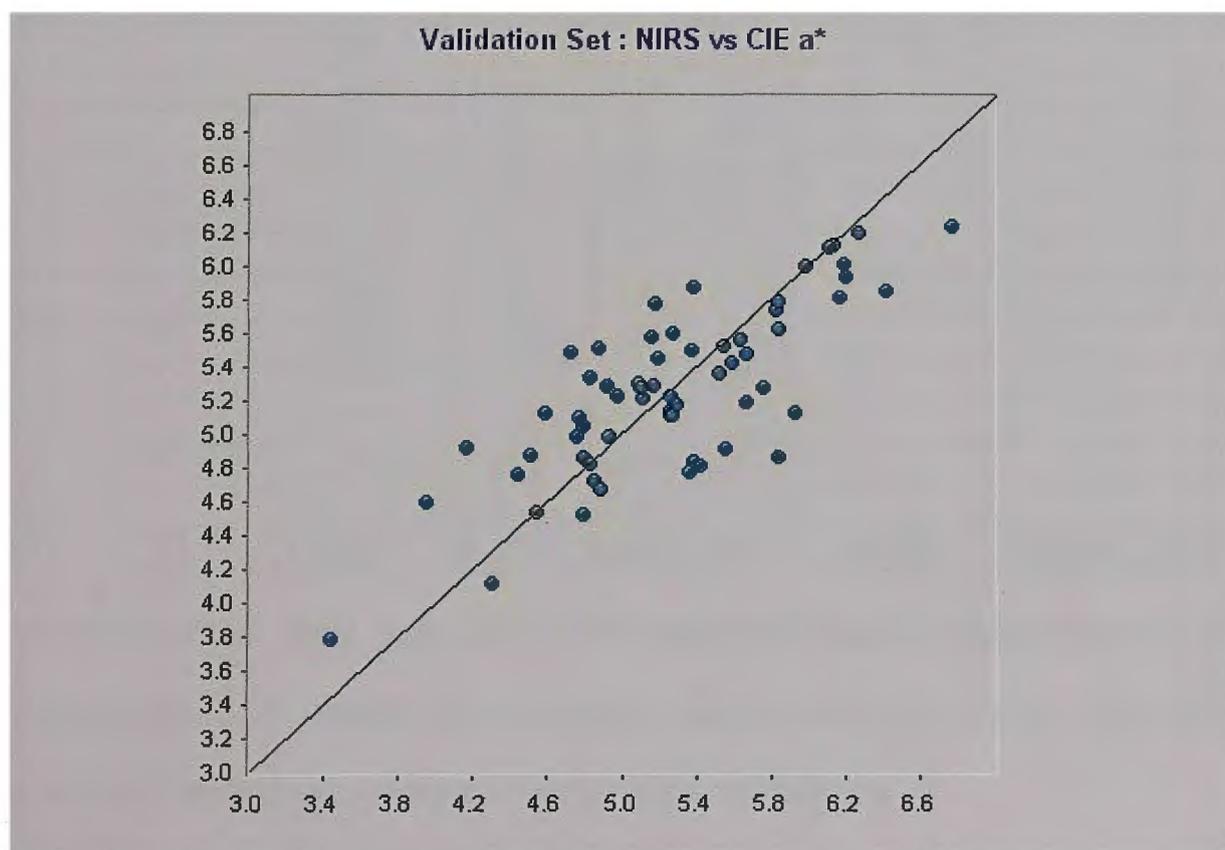


Figure 4.25 The combined seasons validation set for CIE a\* using PLS

Refer to the figure 4.24 for the calibration and figure 4.25 for the validation set of the PLS calibration. Due to the lack of variation in the constituent (the red-green axis of the CIE system) compared to the error inherent in the technique, it was difficult to develop a CIE a\* calibration with a high R<sup>2</sup> value. This was expected as the CIE a\* contribution to sultana colour is minimal. These two calibrations (MLR and PLS) were no exception, in both cases a significant number of outliers were required to be removed to achieve a satisfactory correlation, MLR in particular required the removal of 10% of the calibration sample set which was considered acceptable. For PLS, fewer outliers were removed giving a slightly improved R<sup>2</sup> (Table 4.27 MLR R<sup>2</sup>=0.78; Table 4.28 PLS R<sup>2</sup>=0.79), hence, this calibration would be more stable.

Table 4.28 Factors contributing to the combined PLS calibration of CIE a\* (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

PLS								
Factors	Outliers	R <sup>2</sup> (Cal)	SEC	SECV	F-value	R <sup>2</sup> (Val)	Bias	SEP
1	11	0.243	0.506	0.510	70.23	-	-	-
2	"	0.323	0.480	0.494	52.10	-	-	-
3	"	0.450	0.433	0.463	59.12	-	-	-
4	"	0.491	0.418	0.449	52.16	-	-	-
5	"	0.590	0.376	0.415	61.90	-	-	-
6	"	0.654	0.346	0.402	67.39	-	-	-
7	"	0.667	0.340	0.381	61.06	-	-	-
8	"	0.694	0.327	0.379	60.24	-	-	-
9	"	0.716	0.316	0.373	59.24	-	-	-
10	"	0.746	0.299	0.364	61.77	-	-	-
11	"	0.772	0.284	0.362	64.47	-	-	-
<b>12</b>	<b>11</b>	<b>0.787</b>	<b>0.275</b>	<b>0.355</b>	<b>64.15</b>	<b>0.789</b>	<b>0.0366</b>	<b>0.385</b>

As more factors were included in this calibration steady improvement was observed in both coefficient of determination and SEC (Table 4.28 F=1 R<sup>2</sup>=0.243, SEC=0.506; Factor=12 R<sup>2</sup>=0.787, SEC=0.275)

The overall tendency of the F-value of this calibration was a slight decrease (Table 4.28 F-value=70.2 to 64.2), mirroring the MLR calibration, this showed

that the calibration was a valid model, and could be considered reliable in prediction. The size of the final F-value was relatively large (Table 4.28 F-value=64.15) which also supported this conclusion.

The relatively high difference in the F-value of the MLR and PLS (Table 4.27 MLR F-value=112; Table 4.28 PLS F-value=64.2) calibrations contrasted the results obtained from the validation, however, it was a clear improvement from the PLS calibrations of the individual seasons [15].

The trend seen in the error of cross validation exhibited a steady improvement as factors were added to the calibration (Table 4.28 Factor=1 SECV=0.510; Factor=12 SECV=0.355), without any local minima, which indicated that the calibration would be robust [17]. The validation coefficient of determination achieved by this calibration was 0.789 with an uncorrected SEP of 0.385 and Bias of 0.0366.

## 4.8 Tricolour stimulus b\*

### 4.8.1 Season one CIE b\*

Table 4.29 Laboratory parameters summary of season one b\*

# Samples (in calibration)	# Samples in validation	Range	Mean	Standard Deviation
111 (89)	22	1.01-9.78	6.27	1.915

#### 4.8.1.1 MLR Calibration

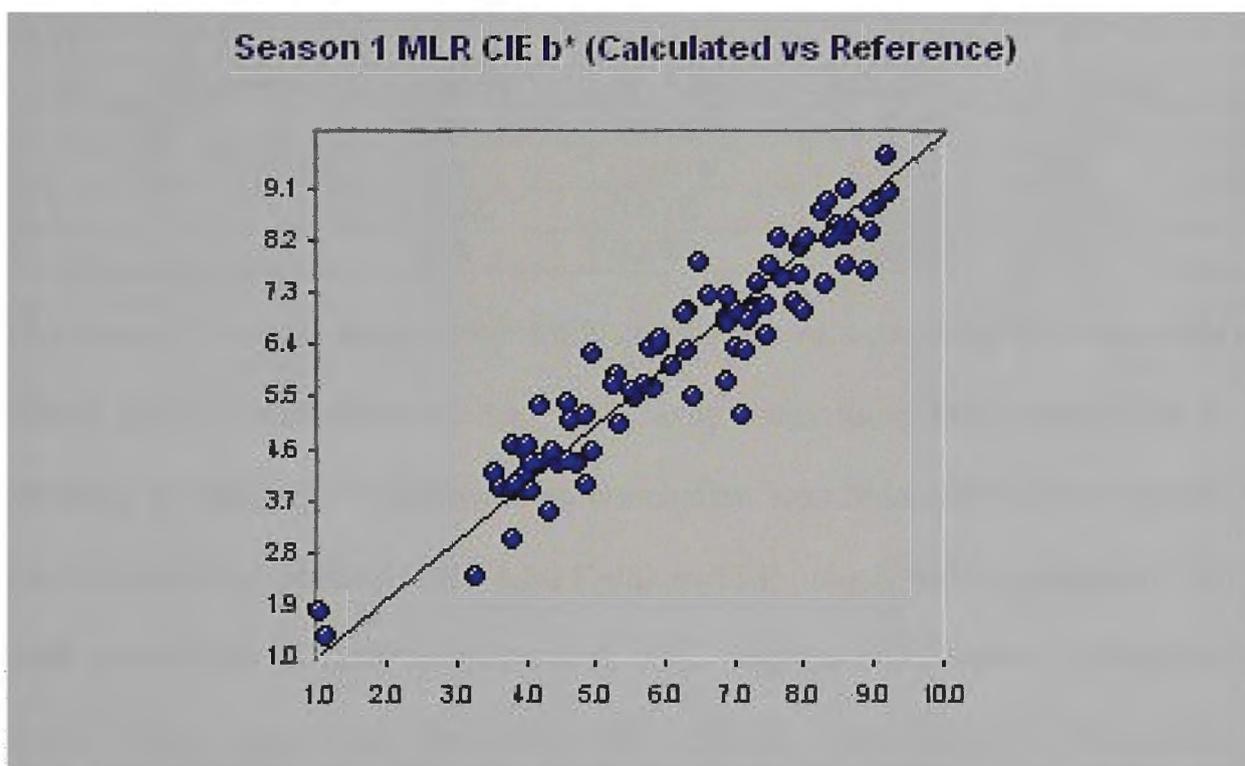


Figure 4.26 A plot of the first seasons calibration for CIE b\* using MLR

Refer above to the plot of the calibration set (Figure 4.26). While the fourth MLR term was located in the visible region of the NIR spectrum [13], the more strongly associated terms were due to features related to the chemical composition of the fruit [12]. The first term (898nm) was attributable to the third overtone stretching vibrations [12] of carbon-hydrogen bonds from any alkyl groups, alkenes with one hydrogen, and aldehydes [14]. The second term (2346nm) arose from a region of the NIR spectra caused by the second overtone of carbon-hydrogen bond deformation vibrations [12], typically seen

when cellulose was present and carbonate ions [13]. The third term (1962nm) was due to the asymmetric stretching vibrations of nitrogen-hydrogen bonds [12], typically seen when amide groups were present within the subject of interest, also tertiary, aryl and carboxylic acid hydroxy groups and water caused vibrations at this wavelength [13]. The fourth term (654nm) was in the visible region of the spectrum [13].

Table 4.30 Wavelengths contributing to the first seasons MLR calibration of CIE b\* (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

MLR					
Term	Outliers	$\lambda$ (nm)	R <sup>2</sup> (Cal)	SEC	F-value
1	-	898	0.8062	0.841	345.3
2	-	2346	0.8476	0.750	228
3	-	1962	0.875	0.684	188.9
4	-	<b>654</b>	<b>0.885</b>	<b>0.660</b>	<b>153.7</b>

The trend in F-values seen in the MLR calibration was one of sharp decrease (Table 4.30 F-value trend=345 to 154), which may have indicated a loss of reliability of prediction, however this conclusion was balanced by the relative size of the final F-value (Table 4.30 F-value=154), which was substantial [15]. Both coefficient of determination and SEC improved as more summation terms were added to the calibration (Table 4.30 Term=1 R<sup>2</sup>=0.806, SEC=0.841; Term=4 R<sup>2</sup>=0.885, SEC=0.660).

### 4.8.1.2 PLS Calibration

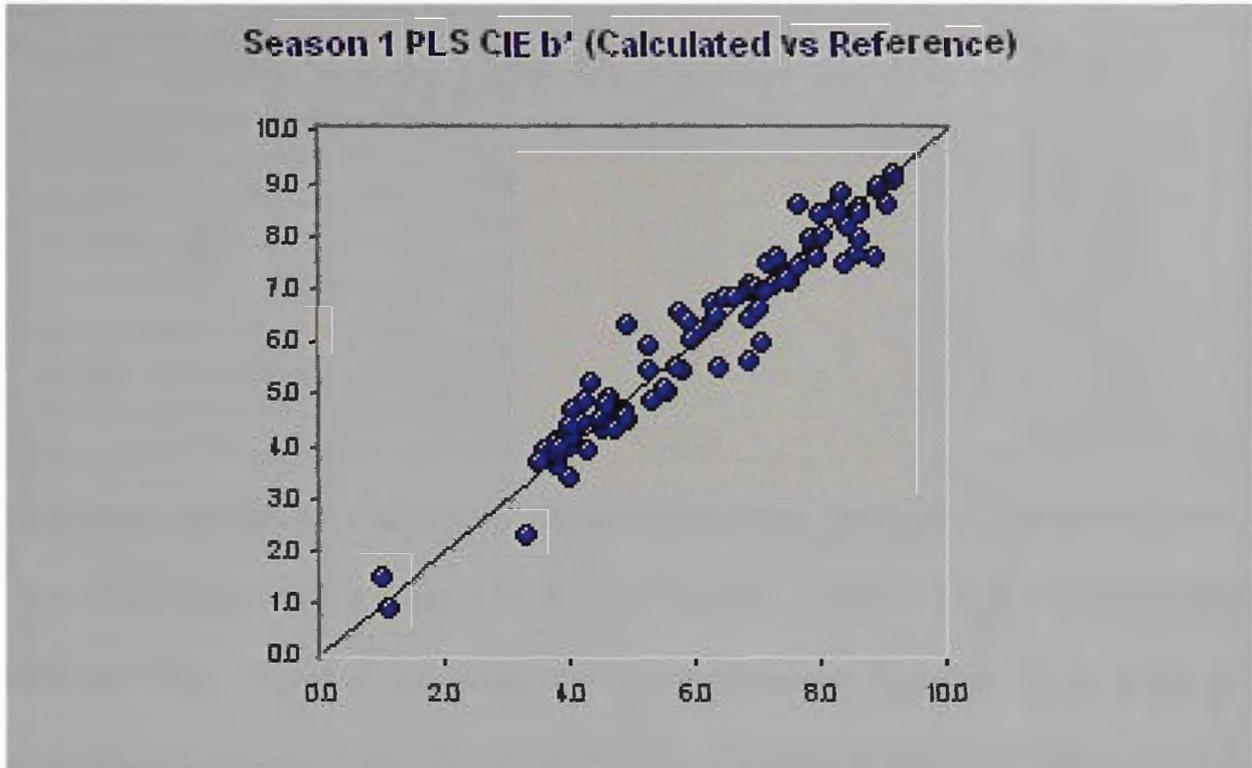


Figure 4.27 A plot of the first seasons calibration for CIE b\* using PLS

See above for the plot of the PLS calibration (Figure 4.27). PLS delivered a more robust calibration than MLR for the first seasons results for CIE b\*, both SEC and the coefficient of determination improved (Table 4.30 MLR  $R^2=0.885$ ,  $SEC=0.660$ ; Table 4.31 PLS  $R^2=0.936$ ,  $SEC=0.509$ ), with minimal loss to the F-value of the calibration. A decreasing trend was shown by the F-value (Table 4.31 F-value trend=139 to 122) of this calibration, however it was not a marked rate of decrease and the final value obtained was relatively large (Table 4.31 F-value=122), both positive indicators of reliability. Contrasting the MLR calibration to the PLS calibration, there is a clear difference between the corresponding F-values of these calibrations (Table 4.30 MLR F-value=153.7; Table 4.31 PLS F-value=121.7) though the discrepancy is not large and the calibration is preliminary, a more precise indication of this calibrations reliability will be found once the seasons have been combined [15].

Table 4.31 Terms used to optimise the first seasons PLS calibration of CIE b\* (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

PLS					
Factors	Outliers	R <sup>2</sup> (Cal)	SEC	SECV	F-value
1	-	0.627	1.17	1.18	139.3
2	-	0.807	0.844	0.931	171.9
3	-	0.866	0.708	0.803	174.4
4	-	0.883	0.667	0.774	150.3
5	-	0.896	0.630	0.762	136.7
6	-	0.908	0.599	0.769	127.8
7	-	0.918	0.569	0.781	122.7
8	-	0.929	0.532	0.809	124.5
9	-	<b>0.936</b>	<b>0.509</b>	<b>0.82</b>	<b>121.7</b>

As further factors were added to the calibration the coefficient of determination and SEC improved (Table 4.31 F=1 R<sup>2</sup>=0.627, SEC=1.17, F=9 R<sup>2</sup>=0.936, SEC=0.509). F-value exhibited an initial increase followed by a gradual decrease (Table 4.31 F=1 F-value=139; F=4 F-value=174; F=9 F-value=122), as the calibration was a preliminary study this was acceptable [15].

The error of cross validation gave a local minimum (Table 4.31 F=1 SECV=1.18; F=5 SECV=0.762) followed by a slight increase as it reached the final factor chosen (Table 4.31 F=9 SECV=0.82), the increase in SECV was slight and, therefore a good preliminary calibration [17]

## 4.8.2 Season two CIE b\*

Table 4.32 Laboratory parameters summary of season two b\*

# Samples (in calibration)	# Samples in validation	Range	Mean	Standard Deviation
95 (76)	19	2.94-9.92	5.74	1.95

### 4.8.2.1 MLR Calibration

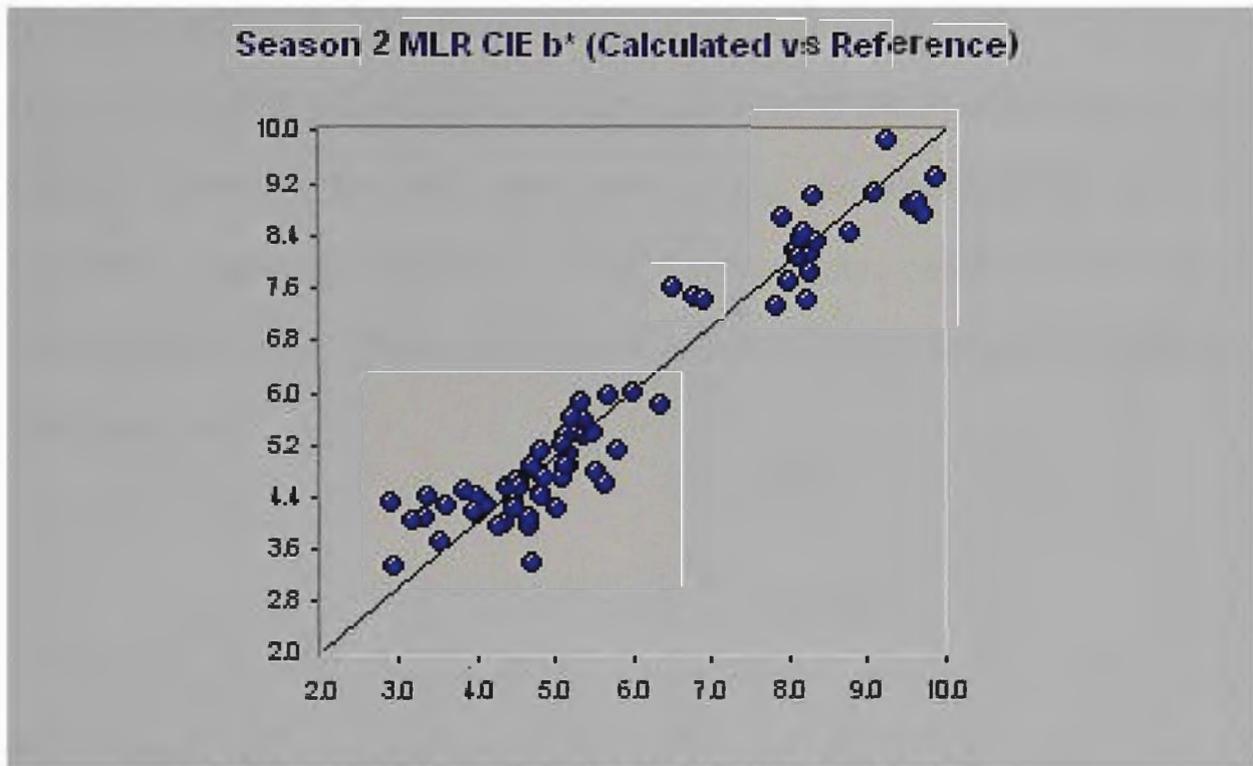


Figure 4.28 A plot of the second seasons calibration for CIE b\* using MLR

Above is provided a plot of the MLR calibration (figure 4.28). The calibration of CIE b\* (Table 4.33), unlike the other parameters taken from measurements using a chromameter, contained a number of terms that originated from the NIR region of the spectra rather than the visible spectrum [13]. The first term (900nm) and the third term (1970nm) displayed the same origins as the wavelengths of the first season calibration for b\* [13]. The second term (1862nm) came from a region that can be attributed to water [13]. The fourth term (518nm) was from the visible region of the spectra [13].

Table 4.33 Wavelengths contributing to the second seasons MLR calibration of CIE b\* (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

MLR					
Term	Outliers	$\lambda$ (nm)	R <sup>2</sup> (Cal)	SEC	F-value
1	-	900	0.850	0.756	384.4
2	-	1862	0.869	0.712	221.8
3	-	1970	0.885	0.670	169.9
<b>4</b>	-	<b>518</b>	<b>0.897</b>	<b>0.639</b>	<b>142.1</b>

As more terms were introduced to the calibration both the coefficient of determination and SEC improved (Table 4.33 Term=1 R<sup>2</sup>=0.85, SEC=0.756; Term=4 R<sup>2</sup>=0.897, SEC=0.639). A sharp decrease in the F-value (Table 4.33 F-value trend=384 to 142) was seen during the development of this calibration, again however, a final F-value was obtained that was comparatively large (Table 4.33 F-value=142) which indicated a reliable calibration [15].

#### 4.8.2.2 PLS Calibration

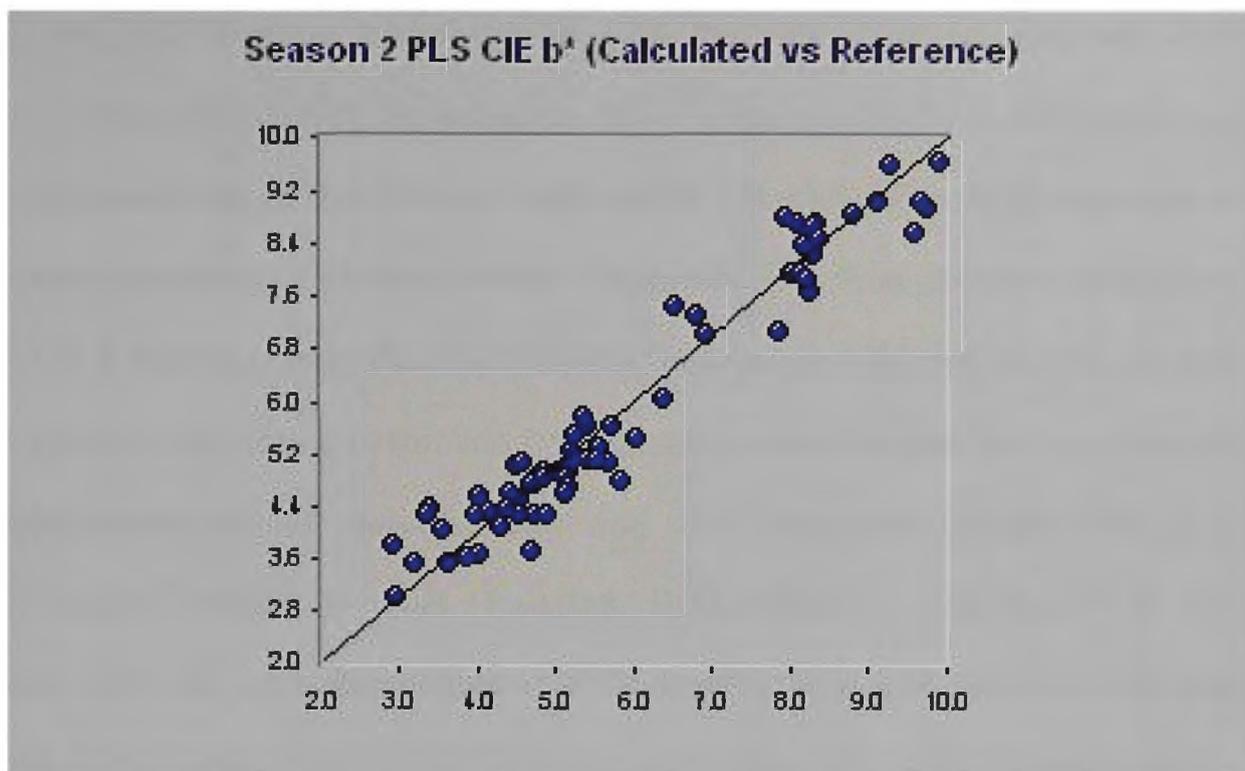


Figure 4.29 A plot of the second seasons calibration for CIE b\* using PLS

Shown above is a plot of the PLS calibration (Figure 4.29). The calibrations developed from the second seasons' results for CIE b\* show a clear performance advantage of PLS over MLR (Table 4.33 MLR  $R^2=0.708$ ,  $SEC=1.05$ ; Table 4.34 PLS  $R^2=0.943$ ,  $SEC=0.491$ ), both the resultant correlation and error of calibration significantly improved, while the corresponding F-value has not been markedly reduced (Table 4.33 MLR F-value=142, Table 4.34 PLS F-value=126).

Table 4.34 Factors contributing to the second seasons PLS calibration of CIE b\* (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

PLS					
Factors	Outliers	$R^2$ (Cal)	SEC	SECV	F-value
1	-	0.708	1.05	1.071	164.5
2	-	0.850	0.761	0.903	189.8
3	-	0.866	0.724	0.795	142.5
4	-	0.898	0.639	0.758	142.4
5	-	0.913	0.594	0.776	133.7
6	-	0.917	0.585	0.747	115.3
7	-	0.931	0.535	0.777	120.4
<b>8</b>	-	<b>0.943</b>	<b>0.491</b>	<b>0.811</b>	<b>126.5</b>

In this calibration, the F-values obtained exhibited a decreasing trend sharper (Table 4.34 F-value trend=164 to 126) than the previous seasons PLS calibration (Table 4.31 first seasons PLS F-value trend=154 to 122), this may have been due to the relative small size of this calibration in comparison to other seasons or different climatic conditions. The final obtained calibration gave a relatively high F-value (Table 4.34 F-value=126), which was a good indication of a sound calibration [15]. As factors were added to the calibration both coefficient of determination and SEC improved (Table 4.34 F=1  $R^2=0.708$ , SEC=1.05; F=8  $R^2=0.943$ , SEC=0.491). The SECV of this calibration showed several minima at the fourth (Table 4.34 SECV=0.758) and the sixth (Table 4.34 SECV=0.747) factors followed by a slight increase as the calibration reached the final factor (Table 4.34 SECV=0.811): a useful preliminary calibration [17].

### 4.8.3 Season three CIE b\*

Table 4.35 Laboratory parameters summary of season three b\*

# Samples (in calibration)	# Samples in validation	Range	Mean	Standard Deviation
108 (86)	22	3.03-9.35	5.49	1.50

#### 4.8.3.1 MLR calibration

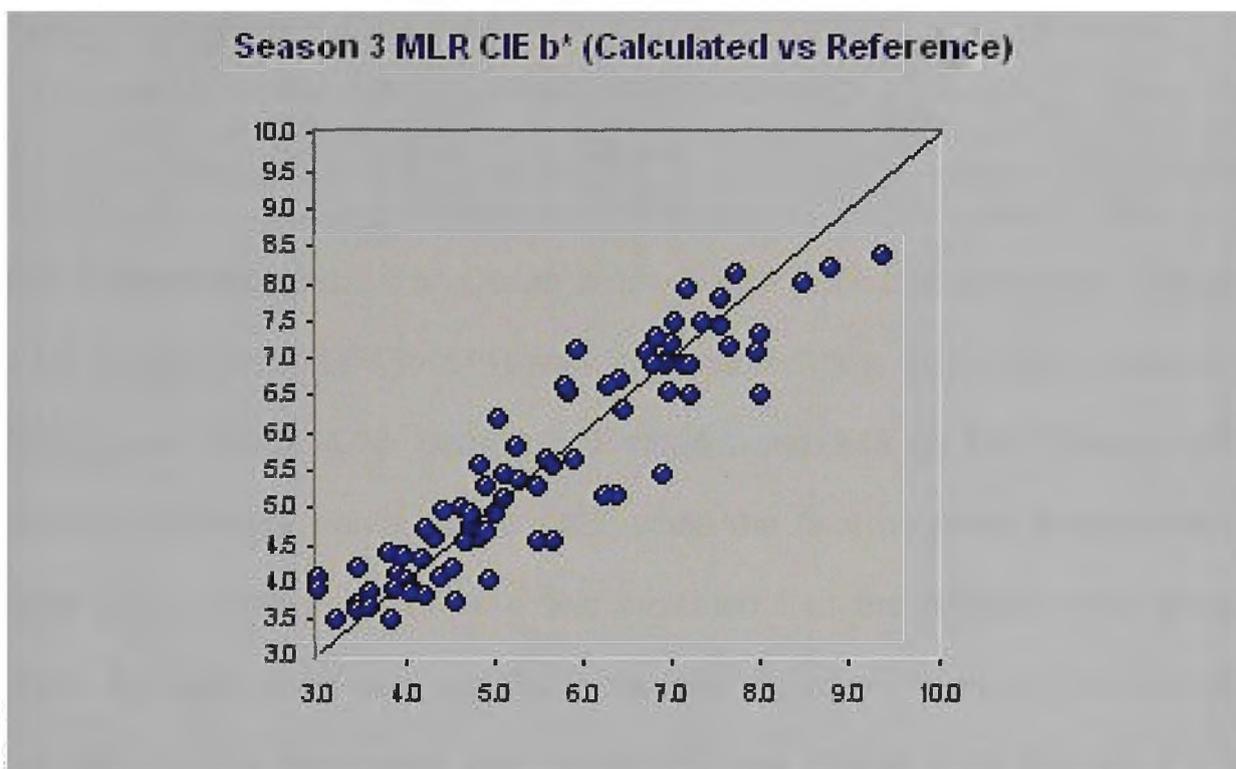


Figure 4.30 A plot of the third seasons calibration for CIE b\* using MLR

CIE b\* originated from the NIR, rather than the visible region of the spectrum. The first term (Table 4.36 900nm) had the same origins as the corresponding terms in the previous seasons calibrations for CIE b\*. The second term (1938nm) was attributed to the oxygen hydrogen bond stretching and deformation vibrations [12] of water molecules, aryl and carboxylic acid, the carbonyl of ketones and esters, primary and tertiary amides, and aldehydes [14]. The third term (514nm), arose from the visible region of the spectrum [13]. The fourth term (1514nm) was linked to the first overtone stretching vibrations of nitrogen hydrogen bonds [12] of amines, primary amides, aqueous ammonia, alkyl and aryl hydroxy groups and water [14]. A plot of the

calibration set (Figure 4.30) is given above. When more terms were included into the calibration there was steady improvement in the coefficient of determination and SEC (Table 4.36 Term=1  $R^2=0.776$ ,  $SEC=0.729$ ; Term=4  $R^2=0.857$ ,  $SEC=0.594$ ).

Table 4.36 Wavelengths contributing to the third season MLR calibration of CIE  $b^*$  (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

MLR					
Term	Outliers	$\lambda$ (nm)	$R^2$ (Cal)	SEC	F-value
1	-	900	0.776	0.729	267.2
2	-	1938	0.802	0.6898	154.3
3	-	514	0.834	0.637	125.4
4	-	<b>1504</b>	<b>0.857</b>	<b>0.594</b>	<b>111.3</b>

The general downward trend seen in the F-values of this calibration (Table 4.36 F-value trend=267 to 111) was less marked than in previous seasons calibrations (Table 4.30 Season 1 F-value trend=345 to 154; Table 4.33 Season 2 F-value trend=384 to 142), while the final obtained F-value was large (Table 4.36 F-value=111), this indicated that the level of over-fitting within the calibration was less than previous seasons, however the overall reliability of the calibration was marginally less (Table 4.30 Season 1 F-value=154; Table 4.33 Season 2 F-value=142) as indicated by the final F-value, which was not surprising as some variation between seasons would be expected and the calibrations were preliminary [15].

### 4.8.3.2 PLS Calibration

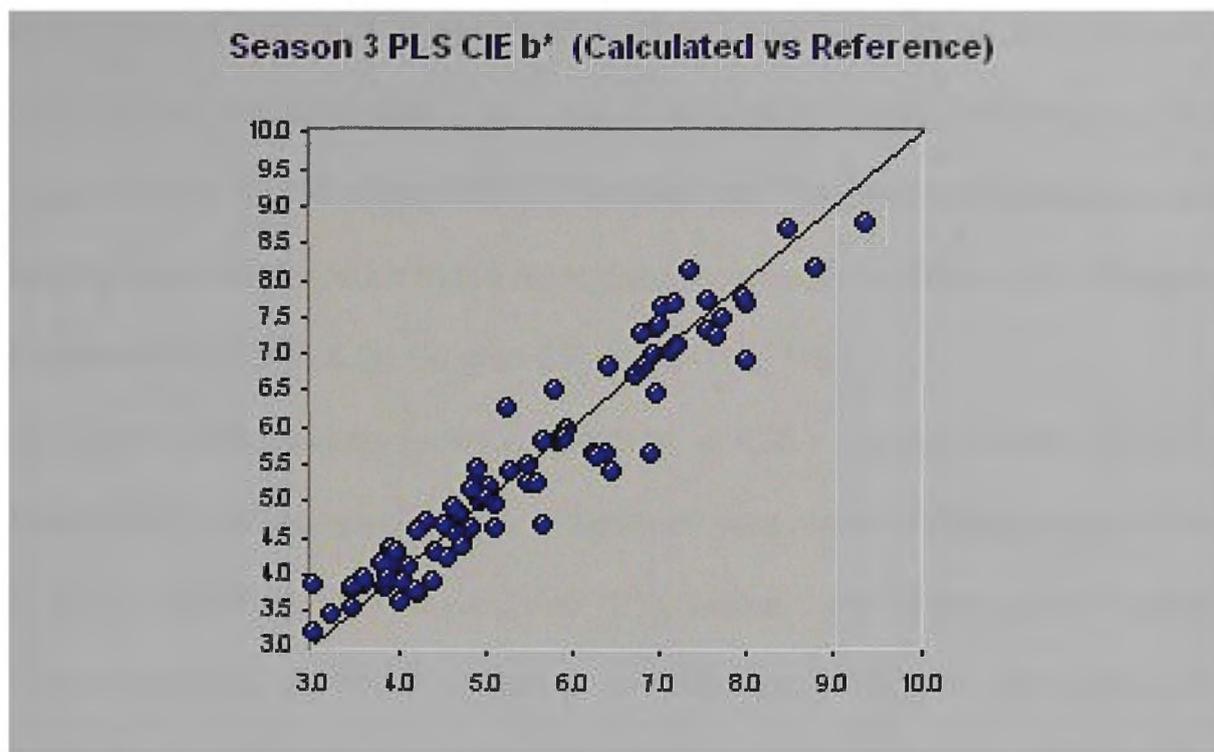


Figure 4.31 A plot of the third seasons calibration for CIE b\* using PLS

A plot of the calibration set is given above (Figure 4.31). The calibrations developed from the third seasons' results reinforced a general trend that PLS typically performed more precisely than MLR (Table 4.37 PLS  $R^2=0.92$ ; Table 4.36 MLR  $R^2=0.86$ ), however the difference was less marked than in previous seasons.

Table 4.37 Factors used to optimise the third season PLS calibration of CIE b\* (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

PLS					
Factors	Outliers	$R^2$ (Cal)	SEC	SECV	F-value
1	-	0.478	1.11	1.13	70.43
2	-	0.651	0.917	0.999	70.76
3	-	0.756	0.771	0.863	77.46
4	-	0.799	0.705	0.788	73.47
5	-	0.849	0.615	0.721	82.28
6	-	0.863	0.590	0.725	75.58
7	-	0.878	0.560	0.776	73.27
8	-	0.904	0.501	0.855	82.25
9	-	<b>0.917</b>	<b>0.47</b>	<b>0.847</b>	<b>84.45</b>

As more factors were included in the calibration a steady improvement was seen in both coefficient of determination and SEC (Table 4.37 F=1  $R^2=0.478$ ,

SEC=1.11; F=9  $R^2=0.917$ , SEC=0.47). The general improvement of the F-value (Table 4.37 F=1 F-value=70.4; F=9 F-value=84.4) of this calibration showed that the calibration itself was satisfactory, but the relatively low final result (Table 4.37 F-value=84.4) indicated that the resultant calibration was slightly less reliable than those from previous seasons (Table 4.31 Season 1 F-value=122; Table 4.34 Season 2 F-value=126) [15].

As seen in the previous years calibrations of CIE  $b^*$  in this project, the trend seen in error of cross validation, a minimum was reached at factor five (Table 4.37 F=1 SECV=1.13; F=5 SECV=0.721), before the final factor was reached (F=9 SECV=0.85), which indicated possible over-fitting in this preliminary calibration, however an overall improvement in the SECV was achieved and such over-fitting that may have occurred would be minimal [17].

#### 4.8.4 Combined Seasons CIE b\*

Table 4.38 Laboratory parameters summary of combined b\*

# Samples (in calibration)	# Samples in validation	Range	Mean	Standard Deviation
311(249)	62	1.01-9.92	5.84	1.82

##### 4.8.4.1 MLR Calibration

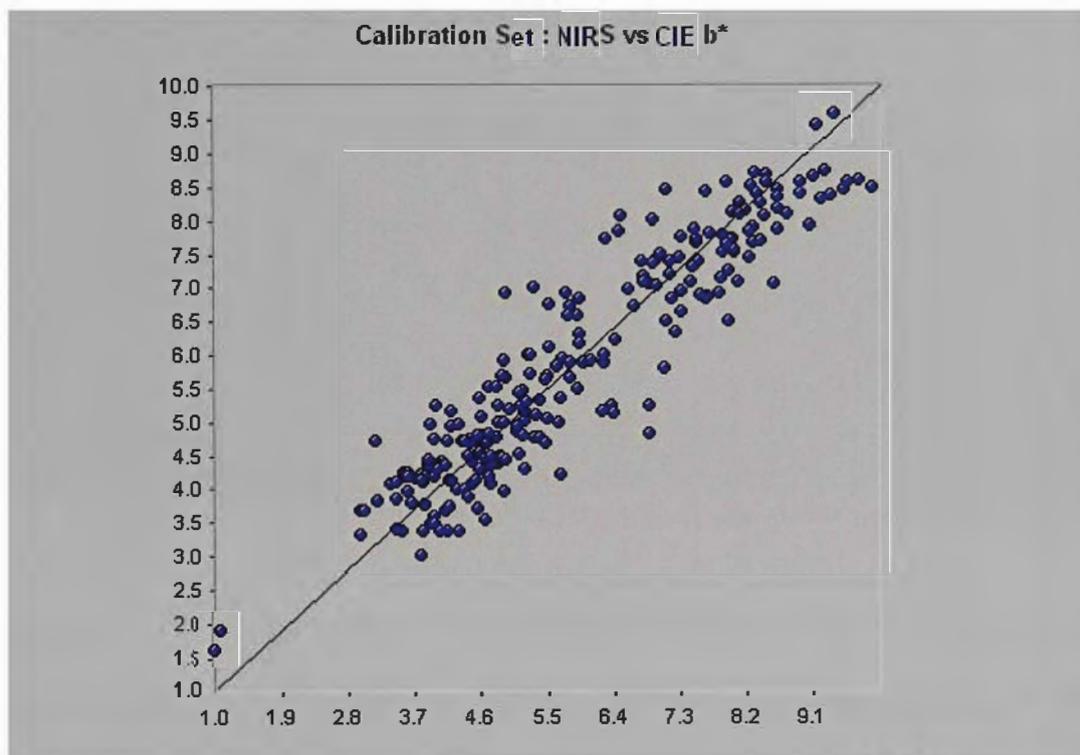


Figure 4.32 A plot of the combined seasons calibration for CIE b\* using MLR

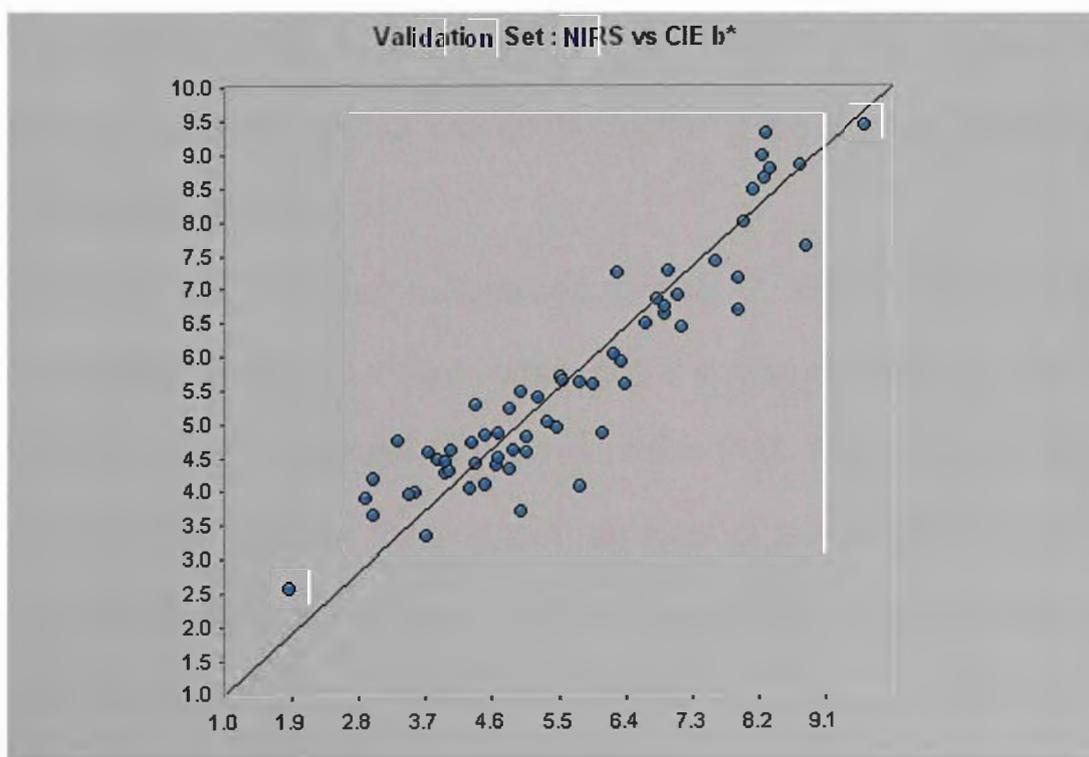


Figure 4.33 A plot of the combined seasons validation for CIE b\* using MLR

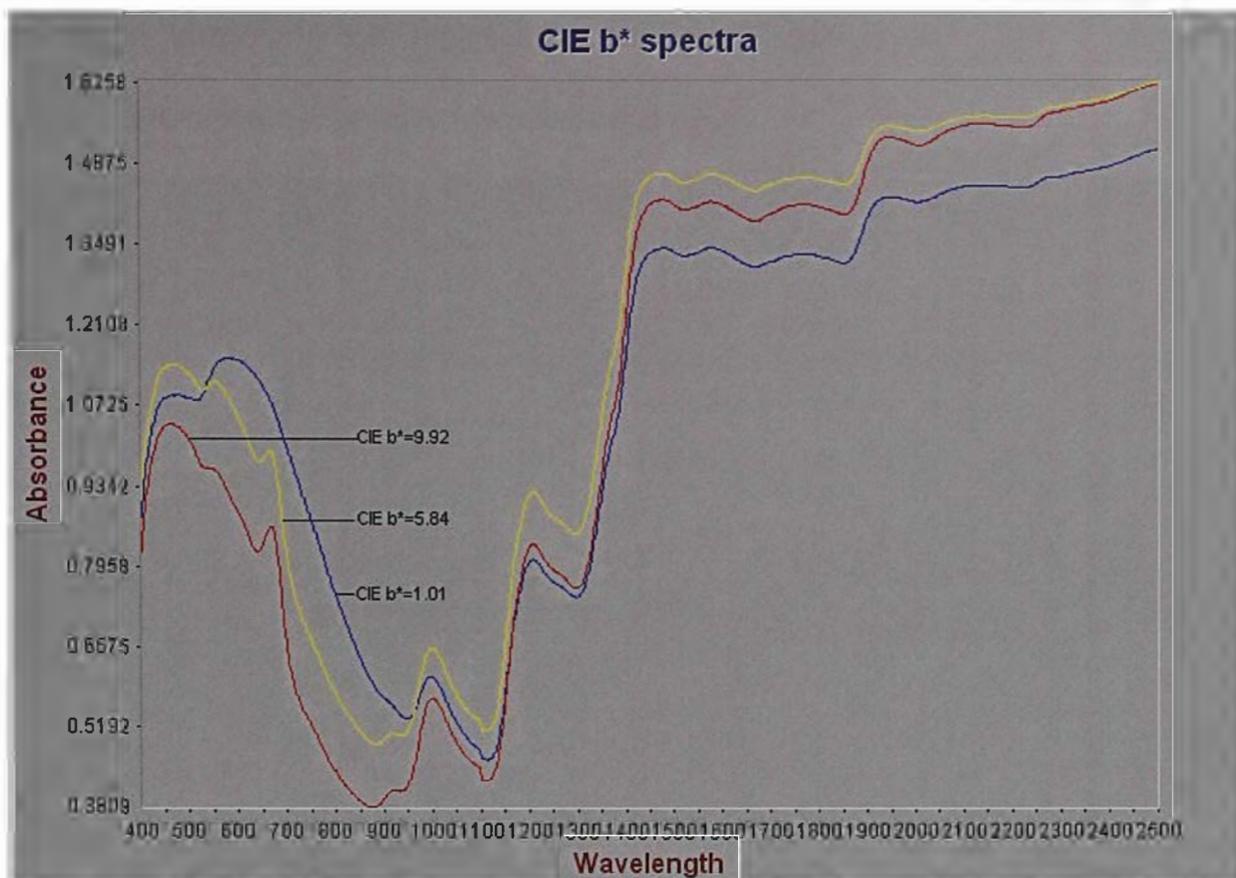


Figure 4.34 Spectra of the minimum (blue), maximum (red) and mean (yellow) samples of CIE  $b^*$

A diagram of three spectra of the calibration set is provided above (Figure 4.34) showing the full spectrum (400 to 2500nm) of the maximum CIE  $b^*$  sample (9.92) the spectrum of the sample closest to the mean value (5.84) and the spectrum of the sample with the lowest CIE  $b^*$  value. Included below is a close up of these spectra around the region of the first wavelength of the MLR calibration (Figure 4.35).

In contrast to the combined calibrations for CIE  $L^*$  and  $a^*$ ,  $b^*$  showed little direct association with the visible region of the spectrum, only the third factor (Table 4.39 604nm) originated from this region [13]. The first term (900nm) was of the same origin as the first term for each of the individual seasons of CIE  $b^*$ . The second term (1936nm) had the same origin as the second term of the third season of CIE  $b^*$ . The fourth term (1674nm) was attributed to the first

overtone stretching vibrations of carbon hydrogen bonds [12] of aromatic compounds, methyl groups and alkenes [14].

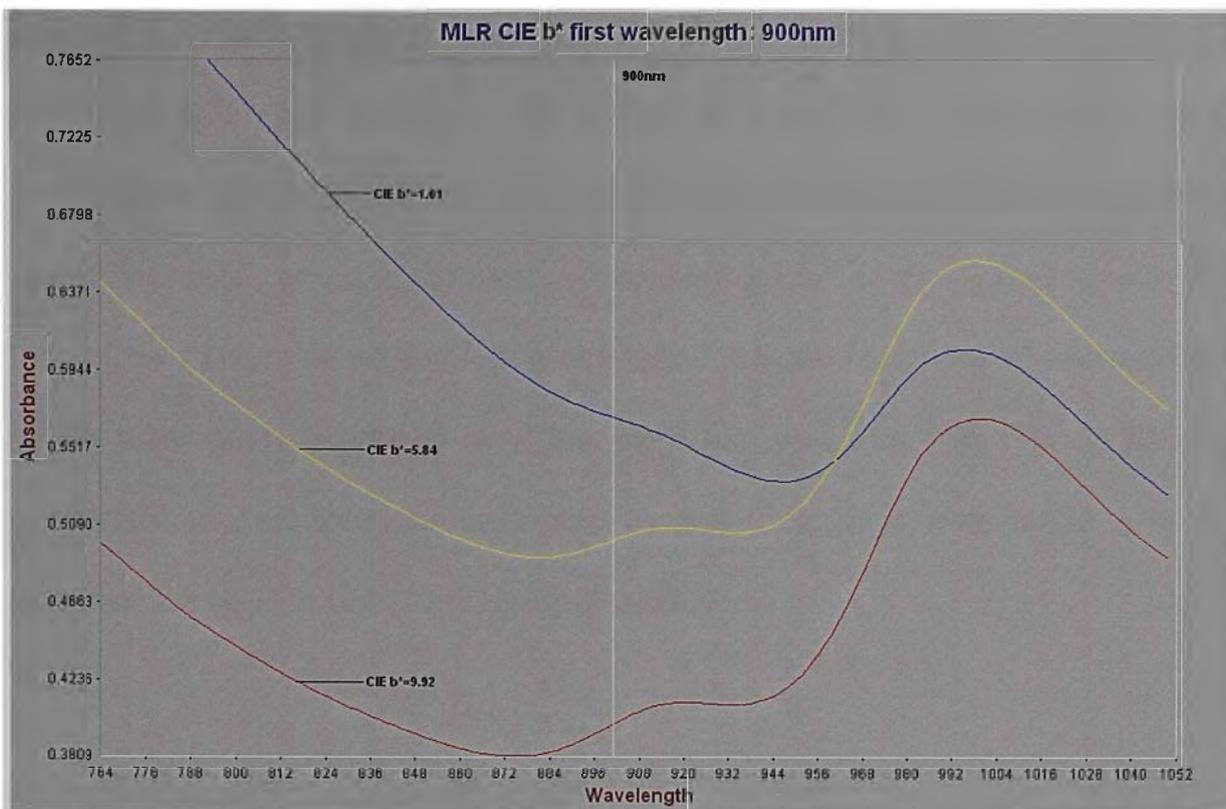


Figure 4.35 Close up of NIR spectra at the first wavelength of the combined MLR calibration. The stretching vibrations of oxygen hydrogen bonds of phenolic compounds [13] were the likely cause of the fifth term (752nm) of this calibration. The sixth term (854nm) arose from the stretching vibrations of carbon-hydrogen and carbon-carbon bonds of aryl compounds [14]. Provided above are plots of the calibration set (Figure 4.32) and the validation set of this MLR calibration (Figure 4.33).

Table 4.39 Wavelengths contributing to the combined MLR calibration of CIE b\* (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

MLR								
Term	Outliers	$\lambda$ (nm)	$R^2$ (Cal)	SEC	F-value	$R^2$ (Val)	Bias	SEP
1	-	900	0.795	0.865	892.7	-	-	-
2	-	1936	0.825	0.764	539.2	-	-	-
3	-	604	0.846	0.717	418.6	-	-	-
4	-	1674	0.855	0.698	335.4	-	-	-
5	-	752	0.863	0.680	284.6	-	-	-
6	-	854	0.869	0.666	249.6	0.932	0.0267	0.677

As more summation terms were added to the calibration to account for variance between the CIE  $b^*$  values and the spectroscopic data, both the coefficient of determination and SEC improved steadily (Table 4.39 Term=1  $R^2=0.795$ ,  $SEC=0.865$ ; Term=6  $R^2=0.869$ ,  $SEC=0.666$ ). The general trend of the F-value of this calibration (Table 4.39 F-value trend=892 to 250) showed a sharp decrease, which is exhibited when the addition of new terms results in a loss of reliability; this was offset by the final F-value (Table 4.39 F-value=250), which was much improved over the individual seasons [15]. The validation coefficient of determination achieved by this calibration was 0.932 with an uncorrected SEP of 0.677 and Bias of 0.0267.

#### 4.8.4.2 PLS Calibration

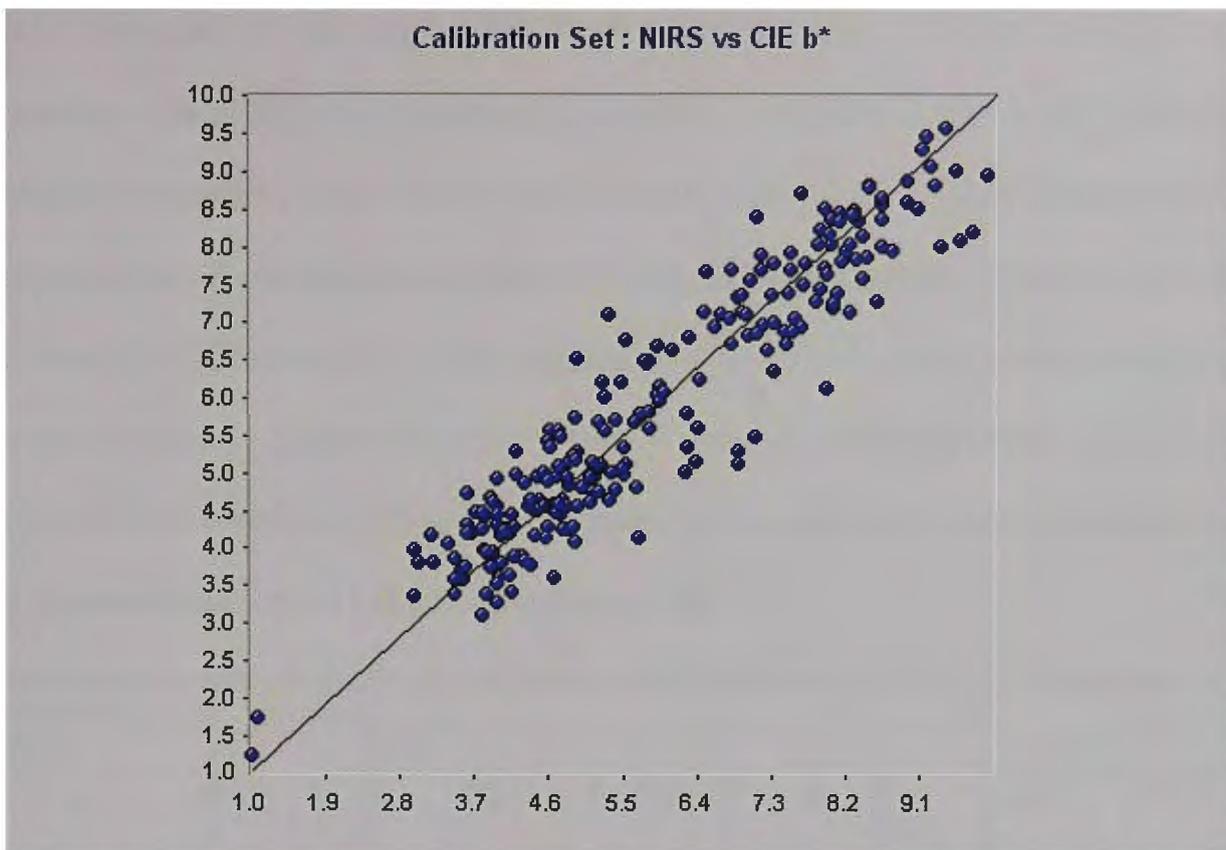


Figure 4.36 A plot of the combined seasons calibration for CIE b\* using PLS

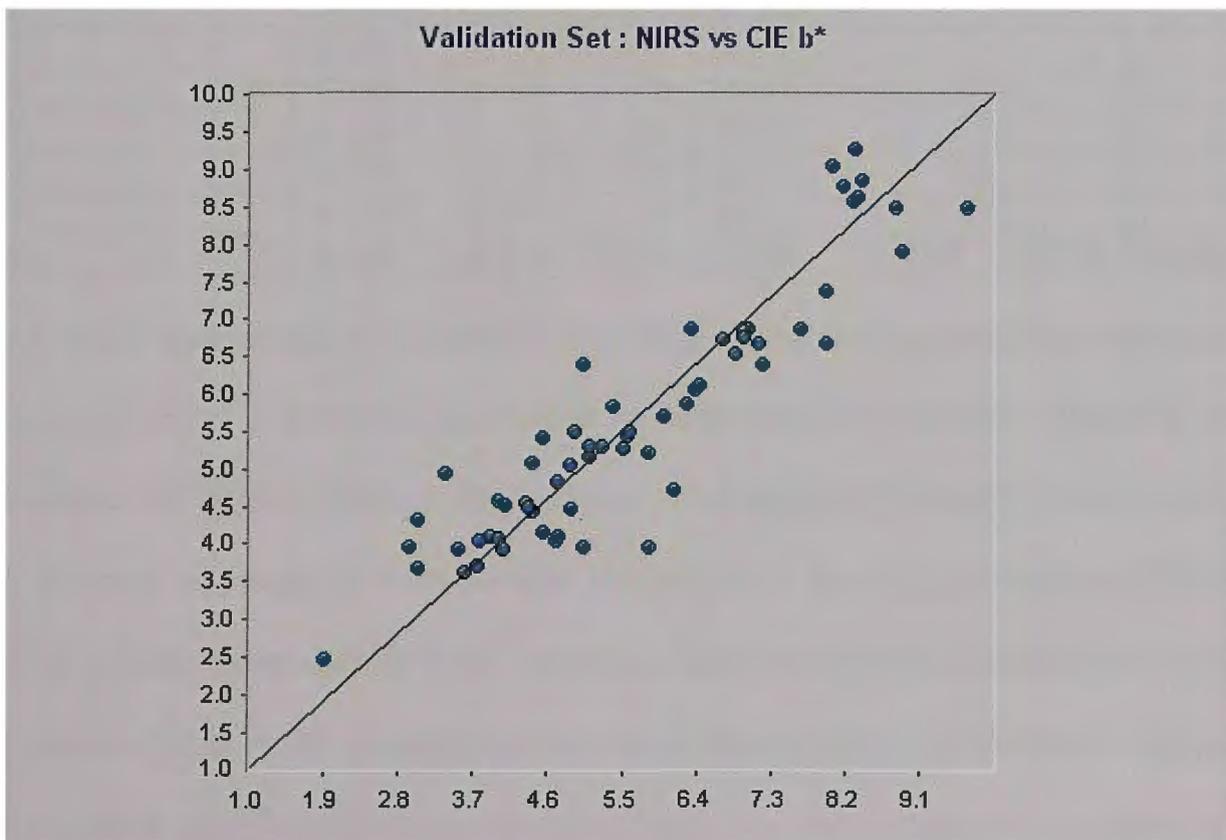


Figure 4.37 The combined seasons validation set for CIE b\* using PLS

Plots of the calibration set (Figure 4.36) and the validation set (Figure 4.37) are given above. CIE b\* in contrast to a\* was the easiest of the CIE

calibrations to develop in both MLR and PLS and it is the most important as it is a measure of the yellowness. Both have provided capable calibrations, however the calibration developed using PLS was more satisfactory than the MLR counterpart (Table 4.40 PLS  $R^2=0.884$ ; Table 4.39 MLR  $R^2=0.869$ ). The difference in F-value between the PLS and MLR calibrations (Table 4.40 PLS F-value=212; Table 4.39 F-value=250 respectively) was slight considering the improvement in performance, and the F-value of the final PLS calibration (Table 4.40 F-value=212) showed a definite improvement when compared to the calibrations of the individual seasons [15].

Table 4.40 Factors contributing to the combined PLS calibration of CIE  $b^*$  (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

PLS								
Factors	Outliers	$R^2$ (Cal)	SEC	SECV	F-value	$R^2$ (Val)	Bias	SEP
1	-	0.614	1.13	1.14	366.3	-	-	-
2	-	0.786	0.844	0.905	421.2	-	-	-
3	-	0.824	0.767	0.806	356.2	-	-	-
4	-	0.839	0.735	0.778	296.5	-	-	-
5	-	0.850	0.712	0.758	255.9	-	-	-
6	-	0.870	0.664	0.728	251.1	-	-	-
7	-	0.878	0.644	0.714	231.4	-	-	-
<b>8</b>	-	<b>0.884</b>	<b>0.630</b>	<b>0.709</b>	<b>212.3</b>	<b>0.924</b>	<b>-0.049</b>	<b>0.646</b>

As more factors were included in the PLS calibration, steady improvement was observed in both the coefficient of determination and SEC (Table 4.40 Factor=1  $R^2=0.614$ , SEC=1.13; Factor=8  $R^2=0.884$ , SEC=0.630). The trend of a general decrease in F-value was continued in the PLS calibration (Table 4.40 F-value trend=366 to 212), however it was far less marked than the MLR F-value (Table 4.39) counterpart and also resulted in a considerably higher final result than the individual seasons. Unlike the trends seen in the individual season calibrations of CIE  $b^*$ , the final calibration exhibited a steady improvement as it reached the final factor (Table 4.40 F=1 SECV=1.14; F=8

SECV=0.71) indicating that the calibration was reliable [17]. The coefficient of determination of the validation set achieved was 0.924 with an uncorrected SEP of 0.646 and Bias of -0.049.

#### **4.9 Comparison to literature examples**

Work by Leroy *et al* [18] used NIR to develop a series of calibrations for the non-destructive determination of the quality of beef, including CIE L\*, a\* and b\* with some success. CIE L\* achieved coefficients of determination of 0.83 for two days *post mortem* meat and 0.85 for eight days *post mortem* beef in the reflectance mode. The calibrations developed from transmission spectra were less successful, giving the determination coefficient of cross validation of 0.68 and 0.64 for two and eight day *post mortem* beef respectively.

The calibration developed for CIE a\* of two and eight day *post mortem* beef achieved a determination coefficient of cross validation of 0.39 and 0.49 for eight days *post mortem* in reflectance mode and for transmission mode gave determination coefficients of cross validation of 0.35 for two days *post mortem* beef and 0.19 for eight days *post mortem* beef.

Calibrations for CIE b\* in reflectance mode yielded determination coefficients of cross validation of 0.75 for two day *post mortem* meat, and 0.73 for eight day *post mortem* meat in the reflectance mode. Calibrations were also developed for CIE b\* in the transmission mode for both two and eight day *post mortem* beef giving determination coefficients of cross validation of 0.56 and 0.44 respectively. These coefficients of determination are not high but the matrix is a difficult one to assess [18].

The work on broiler breast characteristics by Liu and co workers [19] used reflectance NIR spectroscopy to assess, amongst other factors, CIE L\*, a\* and b\*. The calibration for CIE L\* resulted in a determination coefficient of 0.84 and a determination coefficient of the validation set of 0.94. The calibration developed to assess CIE a\* of broiler breast meat achieved a determination coefficient of 0.83, did not to give reliable predictions during validation. Again, this matrix is a difficult product to determine by NIRS.

The calibration for CIE b\* achieved a determination coefficient of 0.78 and validation determination coefficient of 0.80 which considering the nature of the subject achieves a useful performance [19].

NIRS was used by Flinn *et al* [20] to assess CIE tristimulus in whole and ground pulses successfully, calibrations developed from this work for the CIE L\* of ground chickpeas resulted in a determination coefficient of 0.92, with the corresponding whole sample calibration giving a coefficient of determination of 0.96. The calibrations developed for CIE L\* of field peas proved more successful, with a determination coefficient of 0.95 for ground samples and 0.97 for whole samples being obtained.

The calibrations developed to predict CIE a\* gave a determination coefficient of 0.8 for ground Chickpeas and a determination coefficient for ground Field peas of 0.84, whole samples proved more suited to NIRS analysis for colour achieving coefficients of determination for whole chickpeas of 0.94 while whole field peas achieved 0.92. The calibration prepared to assess CIE b\* gave a determination coefficient for whole Chickpeas of 0.96 and 0.92 for whole Field peas. Interestingly, for each case the whole sample calibrations outperformed the ground calibrations, despite this approach being

conceptually more representative [20]. Calibrations developed for the assessment of grains and grain products CIE L\*, a\* and b\* using NIRS was reported by Black and Panozzo [21]. The resultant calibrations of grains and grain products developed from CIE L\* achieved a relatively low coefficient of determination of 0.53 for lentils, while barley yielded a determination coefficient of 0.64 indicating that it would be useable for coarse screening of samples. CIE a\* gave a determination coefficient of 0.865 for flour, 0.94 for whole barley and 0.85 for whole lentils.

CIE b\* calibrations of grain and grain products proved largely successful, being able to predict CIE b\* with a determination coefficient of 0.81 for flour. Whole grains were less successfully assessed achieving determination coefficients of 0.72 for barley and 0.774 for lentils [21].

Chen and Chen [22] developed a series of NIRS calibrations for the assessment of vegetable oil colour (CIE a\* and b\*) using NIR transmittance spectroscopy; a determination coefficient for CIE a\* of 0.99 and a corresponding prediction determination coefficient of 0.98 were obtained. The calibration for yellow (b\*) resulted in a determination coefficient of 0.89 and a corresponding prediction determination coefficient of 0.88 [22].

The variety of agricultural and horticultural products analysed by NIRS to assess CIE tricolour stimulus values gives an indication of the almost universal utility of NIRS to assess quality parameters rapidly. Despite the variety of subjects analysed in the work examined, the differing sample presentation approaches and the different techniques used to develop the calibrations, the results obtained from processed sultanas gives comparable results to those obtained in the research examined.

#### **4.10 Conclusion**

A series of samples of processed sultanas from three seasons from the processors in the Sunraysia Region, of as wide a variety of fruit crown grades and maturities as possible were collected and successfully analysed using a Minolta Chromameter and the CIE colour system.

The data collected were used to develop robust and proven calibrations using both MLR and PLS regression analysis techniques, using NIRS diffuse reflectance spectroscopy. CIE L\* gave a final combined season calibration that achieved a coefficient of determination of 0.799 and a standard error of calibration of 0.780, and gave a prediction coefficient of determination of 0.84. The PLS calibration for CIE L\* yielded a coefficient of determination of 0.835 and a standard error of calibration of 0.727, and achieved a validation coefficient of determination of 0.792.

CIE a\* combined seasons MLR calibration gave a coefficient of determination of 0.777 and a standard error of calibration of 0.262, and yielded a prediction coefficient of determination of 0.794. The PLS calibration for CIE a\* gave a coefficient of determination of 0.787 and a standard error of calibration of 0.275 and resulted in a prediction coefficient of determination of 0.789.

The MLR calibration for CIE b\* yielded a coefficient of determination of 0.869 a standard error of calibration of 0.666 and resulted in a prediction coefficient of determination of 0.932 after testing the calibration against the validation set. The calibration of CIE b\* developed using PLS yielded a coefficient of

determination of 0.884 and a standard error of calibration of 0.63, and gave a prediction coefficient of determination of 0.924.

The rapid assessment of processed samples using NIRS on random batches of samples will begin as soon as the industry is ready to adopt the technique, a limited trial of the calibrations will occur in the coming season.

Addressing the experimental hypothesis:

- This series of calibrations demonstrated successfully that determining colour of processed fruit by NIRS at the processing line would be feasible, and would allow objective assessment of colour to more readily meet customer requirements.

#### 4.11 Bibliography

1. Frank, D. *Investigation of the biochemical basis of browning during the storage of sultanas*, a PhD Thesis. Victoria University, Werribee Campus; 2001.
2. Mark, H. *Principles and practice of spectroscopic calibration*. Brisbane: J Wiley and sons, Inc; 1991; p.31
3. Beebe, KR, Pell, RJ, Seasholtz, MB. *Chemometrics, a practical guide*. Brisbane: J Wiley and Sons, Inc.;1998; p.336.
4. *Near Infrared technology in the agricultural and food industries*. Second Edition. Edited by P Williams and C Norris. St. Paul (Mi): American Association of Cereal Chemists; 2001 p.72-73.
5. *Near Infrared technology in the agricultural and food industries*. Second Edition. Edited by P Williams and C Norris. St. Paul (Mi): American Association of Cereal Chemists; 2001; pp162-163
6. Mark, H. *Principles and practice of spectroscopic calibration*. Brisbane: J Wiley and sons, Inc; 1991; p.107-118.
7. Skoog, D, West, D, Holler, F, Crouch S. *Fundamentals of Analytical Chemistry*, 8<sup>th</sup> Edition. Sydney: Thomson Learning; 2004; p.785-787
8. Sykes, P. *A Guidebook to Mechanism in Organic Chemistry*, 6<sup>th</sup> Edition. Singapore: Longman Ltd; 1996; p.13.
9. Lehninger, A, Nelson, D, Cox, M. *Principles of Biochemistry*, 2<sup>nd</sup> Edition. New York: Worth Publishers; 1993; p.115
10. Miller, JN, Miller, JC. *Statistics and Chemometrics for Analytical Chemistry*. 5th Edition. Sydney: Prentice Hall; 2000; p.124-126
11. Belitz, H-D, Grosch, W, Schieberle, P. *Food Chemistry*. 3<sup>rd</sup> Edition. Heidelberg: Springer; 2004; p.821.

12. The Chart 'Near Infrared Adsorptions: Major analytical bands and their relative peak positions'. Silver Spring (MD): NIRSystems, Inc.; 1992.
13. *Near Infrared technology in the agricultural and food industries*. Second Edition. Edited by P Williams and C Norris. St. Paul (Mi): American Association of Cereal Chemists; 2001. p. 33-36
14. Osborne, B, Fearn, T, Hindle, P. *Near Infrared Spectroscopy with applications in food and beverage analysis*. 2<sup>nd</sup> Edition. Singapore; Longman Scientific and Technical; 1993.
15. *Near Infrared technology in the agricultural and food industries*. Second Edition. Edited by P Williams and C Norris. St. Paul (Mi): American Association of Cereal Chemists; 2001; p.151-152, 180, 194, 196, 212.
16. Beebe, KR, Pell, RJ, Seasholtz, MB. *Chemometrics, a practical guide*. Brisbane: J. Wiley and Sons, Inc.;1998; p335-336
17. *Near Infrared technology in the agricultural and food industries*. Second Edition. Edited by P Williams and C Norris. St. Paul (Mi): American Association of Cereal Chemists; 2001; p.167.
18. Leroy, B, Lambotte, S, Dotreppe, O, Lecocq, H, Istasse, L. "Prediction of Technological and Organoleptic Properties of Beef Longissimus Thorasis from Near Infrared Reflectance and Transmission Spectra", *Meat Science* 2003;66(1):45-54.
19. Liu, Y, Lyon, BG, Windham, WR, Lyon, CE, Savage, EM. "Prediction of Physical Colour and Sensory Characteristics of Broiler Breasts by Visible-NIR Reflectance Spectroscopy", *Poultry Science* 2004;83(8):1467-1474.
20. Flinn, PC, Black, RG, Iyer, L, Brouwer, JB, Meares, C. "Estimating the Food Processing Characteristics of Pulses by Near Infrared

- Spectroscopy, using Ground or Whole Samples”, *Journal of Near Infrared Spectroscopy* 1998;6:213-220.
21. Black, CK, Panozzo, JF. “Accurate Techniques for Measuring Colour Values of Grain and Grain Products using a Visible-NIR Instrument” *Cereal Chemistry* 2004;81(4):469-474.
  22. Y. S. Chen, A. O. Chen. “Quality analysis and purity examination of edible vegetable oils by NIT spectroscopy”. In: Batten, GD, Flinn, PC, Welsh, LA, Blakeney, AB, editors. *Leaping ahead with Near Infrared Spectroscopy*. Melbourne: NIRSG, RACI; 1995; p.316-323.

## **5.0 Results and Discussion II: Quality Assurance of Processed Sultanas – Water Activity, Protein, Titratable Acidity and Lipids.**

### **5.1 Introduction**

Owing to the fact that increased packaging labelling requirements could be introduced in the future, the ability to easily include more detailed information, by rapidly assessing the processed fruit with a non-invasive technique was given high priority. The near infrared spectrophotometer was considered the most suitable instrument for such analysis, as NIRS has a broad range of industrial and agricultural applications, is non-destructive, non-intrusive and rapid.

Following on from previous NIR projects at Victoria University for the Dried Fruit Industry that met key industry requirements for the rapid assessment of protein, water activity, titratable acidity, on unprocessed fruit, a series of analyses on processed sultanas was proposed to give further benefits to the Industry as outlined in the following section.

Firstly, a rapid means of acquiring feedback to the fruit processors, for example water activity by NIRS, a parameter linked to the total water content was considered important, this parameter would give an indication of the storage instability of fruit following processing, as further deterioration due to browning is possible [1,2]. Secondly, valuable information can also be obtained in regards to the amount of washing and drying the fruit require during processing, a parameter currently done by a conductivity measurement which is slow and time consuming.

Protein would have two possible applications. Firstly, it would give a strong indication of browning instability post processing and secondly it is a common nutritional parameter for processed foods.

Lipids analyses were also considered in order to determine the amount of finishing oil used on the fruit during processing, an ingredient added to improve the flow properties of the fruit by retarding the fruits natural adhesiveness, which often causes clumping [3]. Titratable acidity another parameter selected for analysis as it gives an indication of the fruits maturity and “full bodied-ness”, which is an important quality parameter largely overlooked during classical assessments [4].

The main intention of the following chapter was as follows.

After collecting samples from three seasons from the processors in the Sunraysia Region, of as wide a variety of fruit crown grades and maturities as possible. This series of samples were then analysed using Kjeldahl nitrogen, water activity, percent lipids and titratable acidity. This data were then used to develop robust and proven calibrations using both MLR and PLS regression analysis techniques, by means of NIRS diffuse reflectance spectroscopy. Pre-treatments for the calibrations in all cases were N-point smooth and 2<sup>nd</sup> derivative. Following the successful development of these calibrations the rapid assessment of processed samples using NIRS on random batches of samples was then trailed in conjunction with the Sunraysia dried fruit industry to demonstrate the methods validity and ease of operation.

Experimental hypothesis:

- The successful development of calibrations capable of determining quality parameters of sultanas currently not assessed in processed fruit, thus allowing objective assessment of quality, pro-actively preparing the sultana processing industry for possible changes in processed food labelling regulations.

## 5.2 Water activity

### 5.2.1 Season one water activity

Table 5.1 Laboratory parameters summary of season one Aw

# Samples (in calibration)	# Samples in validation	Range	Mean	Standard Deviation
111 (89)	22	0.435-0.604	0.547	0.025

#### 5.2.1.1 MLR calibration

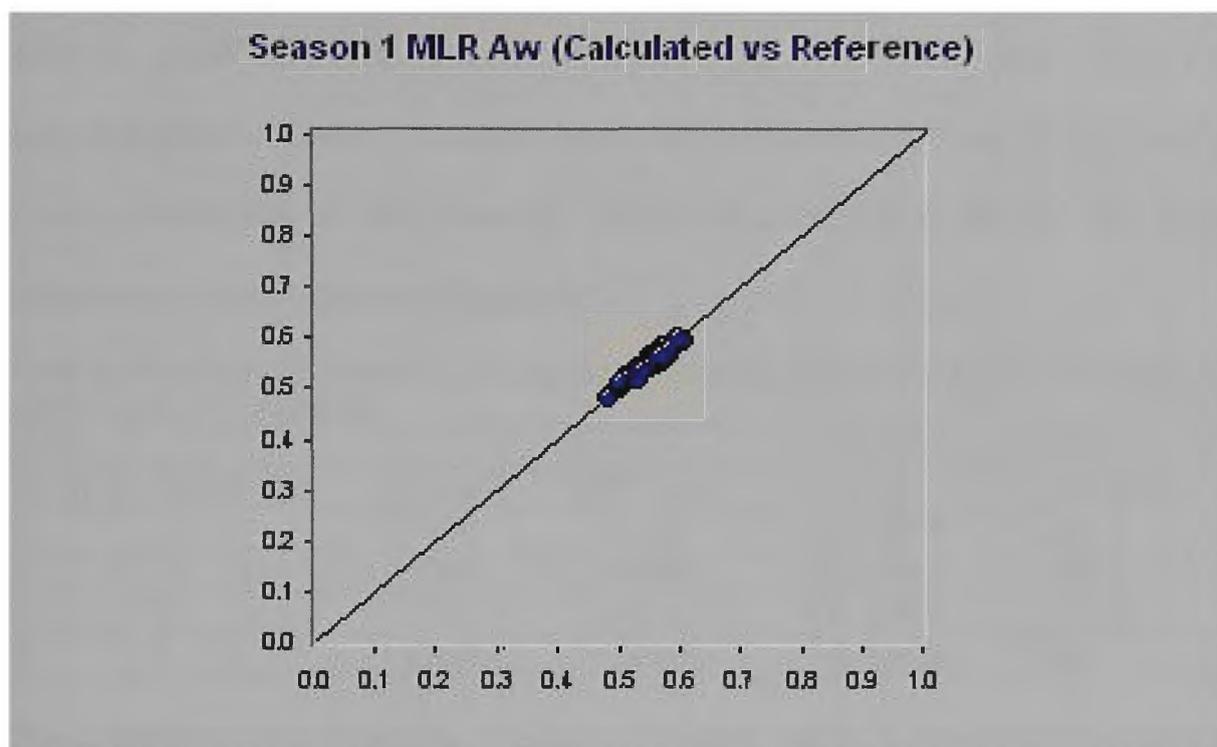


Figure 5.1 A plot of the first seasons' calibration of water activity using MLR.

The following assignments are theoretical and as such there are a number of possible origins consistent with the complex matrix of the medium being studied. Further investigation beyond the scope of this project would need to be completed to confirm these assignments. Refer to section 4.6 for further details. The terms used to develop the calibration for predicting water activity from the first seasons samples, were all from the NIRS region of the spectrum, however the first term (Table 5.2 1276nm) was in the second overtone region, close to the 1<sup>st</sup> overtone band of C-H combinations [5]. The

second term (1064nm) was attributed to a second overtone nitrogen-hydrogen stretching vibration typically seen when amines are present within the subject of interest [5], also amides, alkyl groups, aldehydes, as well as alkyl, carboxylic acid and aryl hydroxy groups [6]. The third term (1354nm) was identified as due to carbon-hydrogen vibrations [5] caused by both stretching and deformation due to the presence of primary and tertiary alkyl groups, alkenes, aldehydes, and secondary amides [6]. The fourth term (2356nm) was indicative of second overtone deformation vibrations in carbon-hydrogen bonds of alkanes in cellulose [7]. Above is provided a plot of the MLR calibration of water activity (Figure 5.1).

Table 5.2 Wavelengths contributing to the MLR calibration of water activity (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

MLR					
Term	Outliers	$\lambda$ (nm)	$R^2$ (Cal)	SEC	F-value
1	-	1276	0.906	0.0079	790.4
2	-	1064	0.922	0.0072	478.3
3	-	1354	0.932	0.0068	367.6
<b>4</b>	-	<b>2356</b>	<b>0.939</b>	<b>0.0065</b>	<b>304</b>

The calibration developed by adding summation terms to account for variance between the laboratory and spectroscopic data, both the coefficient of determination and SEC improved (Table 5.2 Term=1  $R^2=0.906$ , SEC=0.0079; Term=4  $R^2=0.939$ , SEC=0.0065), while the F-value of the calibration decreased as more terms were introduced (Table 5.2 Term=1 F-value=790.4 Term=4 F-value=304) the final value obtained was large indicating that the calibration would give reliable predictions [8].

### 5.2.1.2 PLS calibration

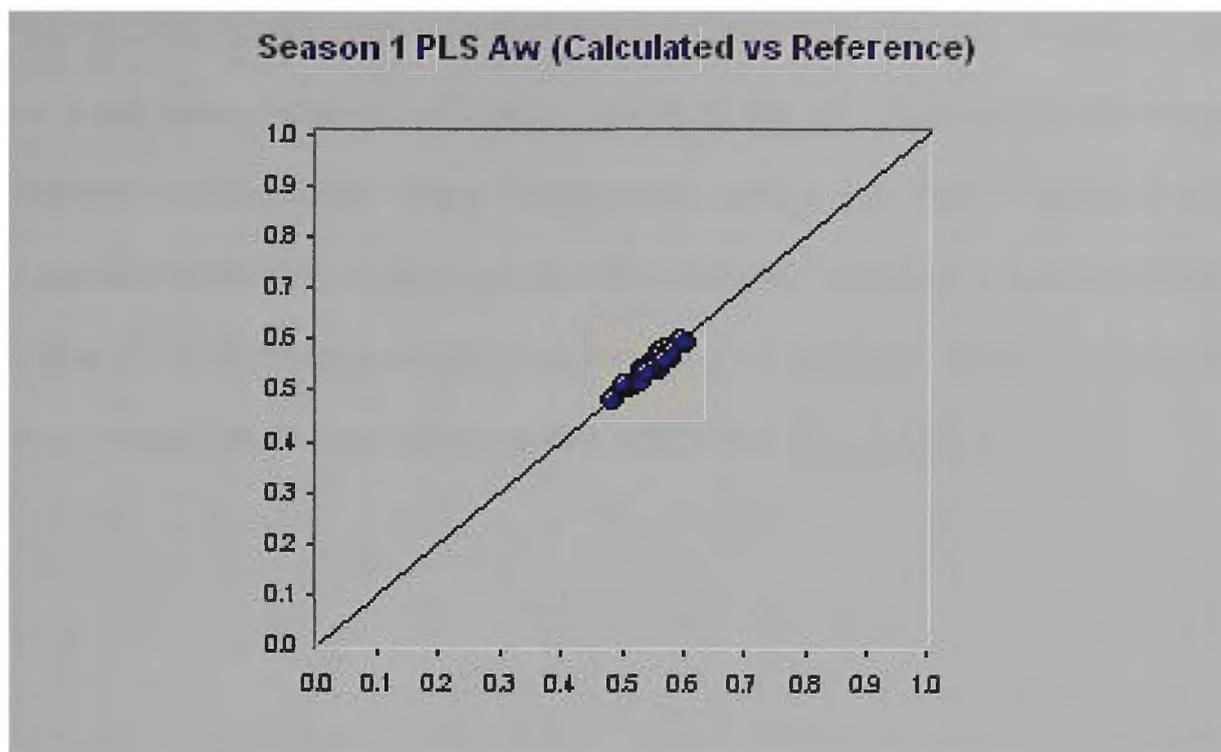


Figure 5.2 A plot of the first seasons' calibration of water activity using PLS

Refer to the diagram for a plot of the PLS calibration of water activity (Figure 5.2). From calibrations developed from the first seasons' water activity results, PLS showed increased robustness when compared to MLR. Both the correlation and error of calibration were higher when compared to the MLR, respectively (Table 5.2 MLR  $R^2=0.939$ ,  $SEC=0.0065$ ; Table 5.3 PLS  $R^2=0.941$ ,  $SEC=0.0064$ ).

Table 5.3 Factors contributing to the PLS first seasons' calibration of water activity (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

PLS					
Factors	Outliers	$R^2$ (Cal)	SEC	SECV	F-value
1	-	0.333	0.021	0.023	40.99
2	-	0.668	0.149	0.0161	81.57
3	-	0.906	0.0079	0.0093	258.60
4	-	0.928	0.0070	0.0082	254.58
5	-	<b>0.941</b>	<b>0.0064</b>	<b>0.0079</b>	<b>253.42</b>

A reduction of the F-value between PLS and MLR was seen: but it was minimal (Table 5.2 MLR F-value=304; Table 5.3 PLS F-value=253). The trend shown in the standard error of cross validation was one of steady

improvement as the calibration was developed (Table 5.3  $F=1$   $SECV=0.023$ ;  $F=5$   $SECV=0.0079$ ), and indicated that the calibration was not over-fitted, as the predictions made during cross validation are not deteriorating from the reference results when more factors were added [9]. The F-value of the calibration sharply increased as more factors were included in the calibration (Table 5.2  $F=1$   $F\text{-value}=40.99$ ;  $F=5$   $F\text{-value}=253.42$ ) and resulted in a large final F-value, which was indicative of a valid predictive model [8].

## 5.2.2 Season two water activity

Table 5.4 Laboratory parameters summary of season two Aw

# Samples (in calibration)	# Samples in validation	Range	Mean	Standard Deviation
95 (76)	19	0.4795-0.607	0.538	0.023

### 5.2.2.1 MLR Calibration

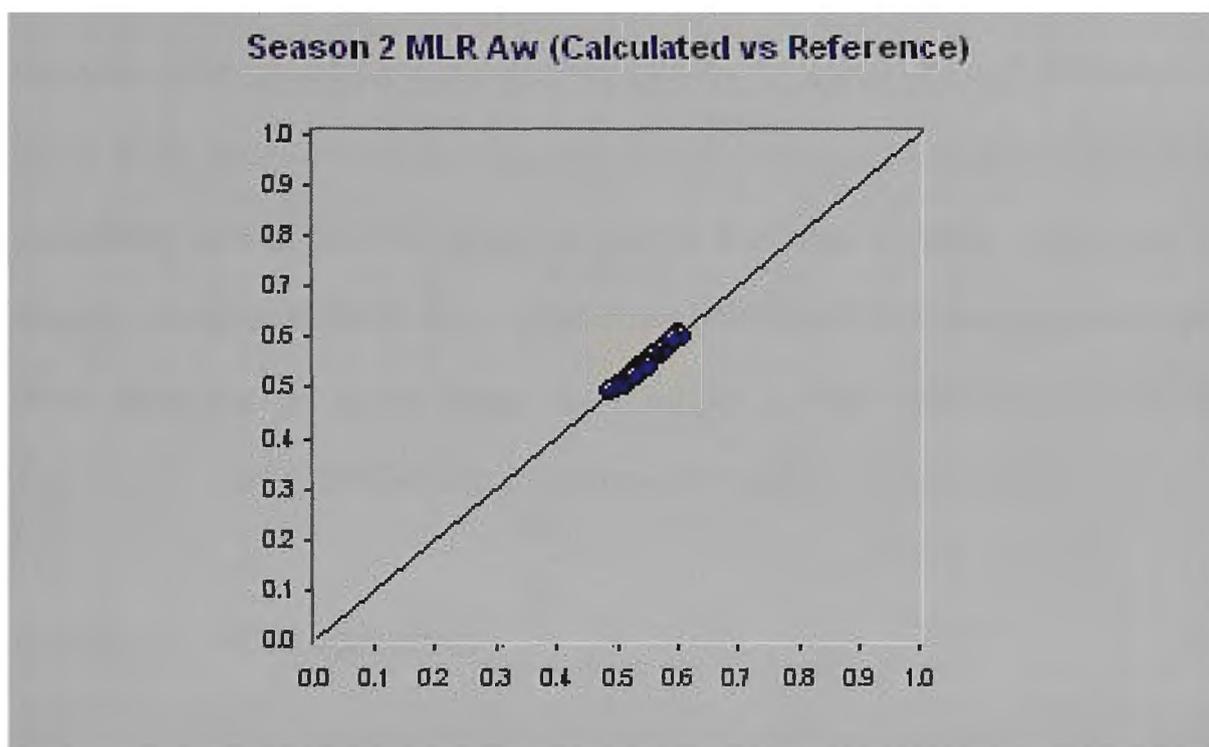


Figure 5.3 A plot of the second seasons' calibration for water activity using MLR

The first term (Table 5.5 1276nm) originated from the same region as the previous season MLR calibration for Aw [5]. The second term (496nm) in this calibration of water activity by dew point, came from the visible region of the spectra [7]. The third term (844nm) was linked to carbon-carbon and carbon hydrogen stretching vibrations of aromatic functionality [7]. The fourth term (1132nm) was linked to the second overtone of stretching vibrations [5] of carbon-hydrogen bonds in aryl groups methyl and tertiary alkyl, alkenyl, aryl and aldehyde groups, as well as primary and secondary amides [6]. Given above is a plot of the calibration (Figure 5.3). A general decrease in the F-

Table 5.5 Wavelengths contributing to the second seasons' MLR calibration of water activity (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

MLR					
Term	Outliers	$\lambda$ (nm)	$R^2$ (Cal)	SEC	F-value
1	-	1276	0.931	0.0059	917.8
2	-	496	0.951	0.005	645.3
3	-	844	0.966	0.0042	624.4
4	-	1132	0.971	0.0039	541.5

value was seen during the development of this calibration, which implied that the calibration would be more reliable with fewer terms (Table 5.5 Term=1 F-value=918; Term=4 F-value=542), the final F-value however indicated that the calibration would provide valid results as the final F-value was large [8]. Steady improvements in both correlation coefficient and error of calibration were observed as more terms were added to this calibration (Table 5.5 Term=1  $R^2=0.931$ , SEC=0.0059; Term=4  $R^2=0.971$ , SEC=0.0039).

### 5.2.2.2 PLS calibration

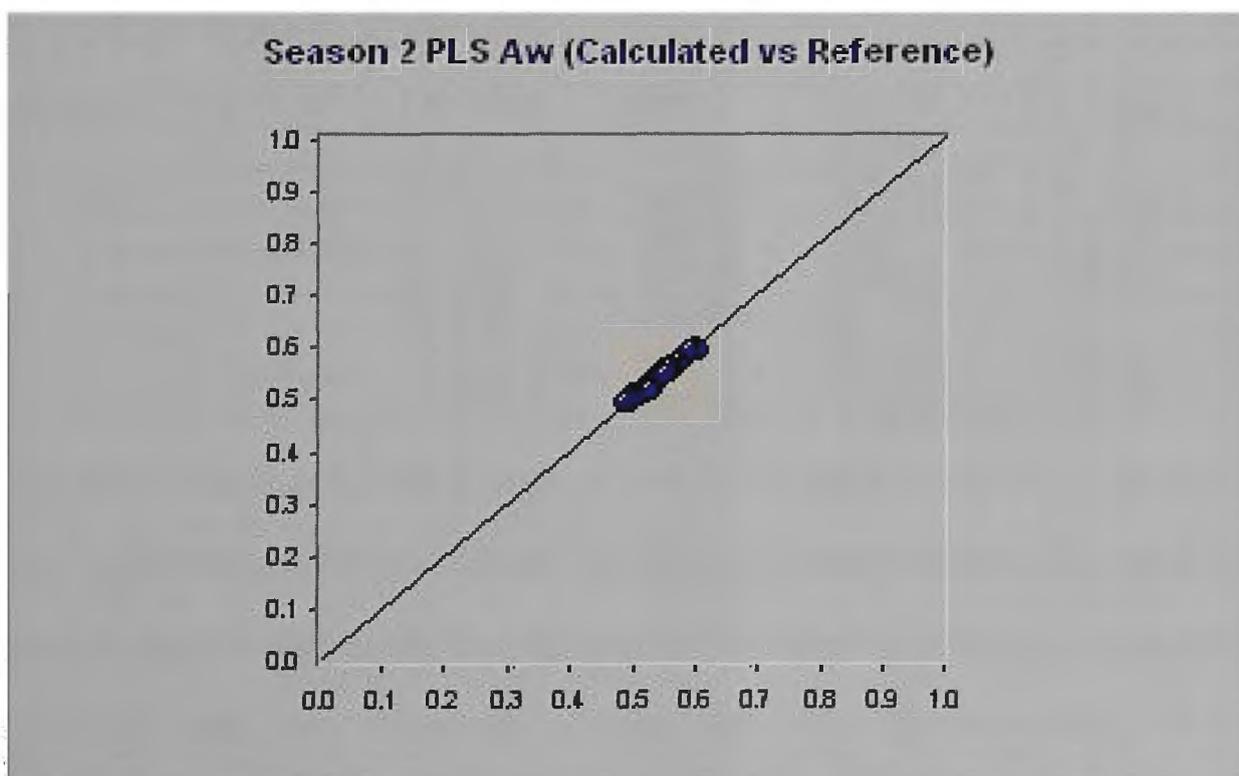


Figure 5.4 A plot of the second seasons' calibration for water activity using PLS

Provided above is a plot of the PLS calibration (Figure 5.4). The calibrations developed from the results obtained over the course of the second season, have reversed the trends observed in previous examples, the correlation and SEC of the MLR (Table 5.5  $R^2=0.971$ ,  $SEC=0.0039$ ) calibration is higher compared to the PLS calibration (Table 5.6  $R^2=0.964$ ,  $SEC=0.0044$ ), the F-value difference was large also, as the PLS F-value was markedly smaller than the MLR value, which may affect prediction (Table 5.5 MLR F-value=541; Table 5.6 PLS F-value=239.01), however the data were from only one season hence a more valid indicator of the calibrations capabilities will be seen when the seasons are combined. A strong increase in the F-value was seen during the development of this calibration (Table 5.6 F=1 F-value=12.24; F=7 F-value=239.01), this showed that the calibration was not over-fitted by the introduction of factors [8].

Table 5.6 Factors contributing to the second seasons' PLS calibration of water activity (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

PLS					
Factors	Outliers	$R^2$ (Cal)	SEC	SECV	F-value
1	-	0.152	0.0206	0.0208	12.24
2	-	0.636	0.0136	0.0154	58.53
3	-	0.887	0.0076	0.0093	173.56
4	-	0.906	0.0070	0.0085	156.86
5	-	0.949	0.0052	0.0071	237.69
6	-	0.959	0.0047	0.0068	244.21
7	-	<b>0.964</b>	<b>0.0044</b>	<b>0.0067</b>	<b>239.01</b>

The final F-value obtained is large (Table 5.6 F-value=239), but is smaller than MLR (Table 5.5 F-value=542). The error of cross validation displayed a gradual decrease indicating that the prediction achieved with cross validation improved with the calibration (Table 5.6 F=1 SECV=0.0208; F=7 SECV=0.0067), which was an indication that the calibration was not over-fitted [9].

### 5.2.3 Season three water activity

Table 5.7 Laboratory parameters summary of season three Aw

# Samples (in calibration)	# Samples in validation	Range	Mean	Standard Deviation
108 (86)	22	0.494-0.612	0.545	0.022

#### 5.2.3.1 MLR calibration

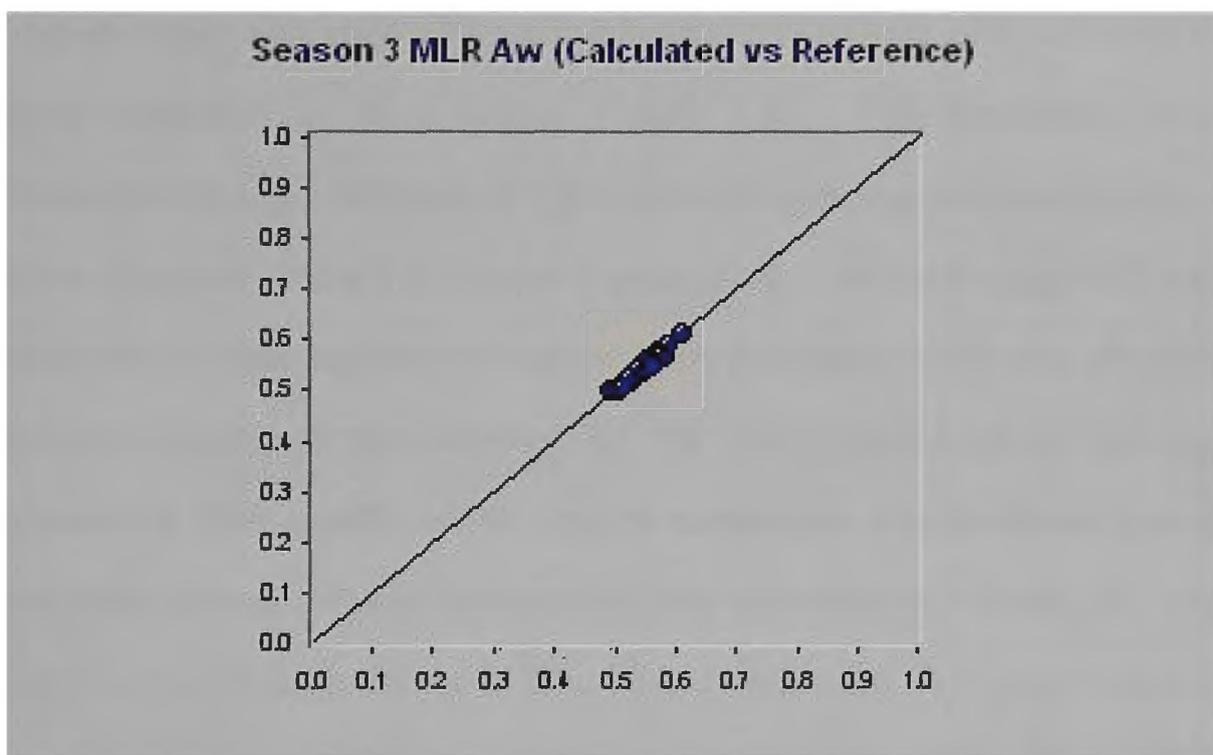


Figure 5.5 A plot of the third seasons' calibration for water activity using MLR

The first term (Table 5.8 1276nm) of the water activity calibration was the same as seen in seasons one and two (Tables 5.2 and 5.5). The second term (1224nm) was due to the second overtone stretching vibration [5] of carbon hydrogen bonds of alkyl groups and alkenyl groups with one hydrogen bonded [6]. The third term (1064nm) was linked to second overtone stretching vibrations of nitrogen hydrogen bonds [5] of amines, also the vibrations of hydroxy groups, amides and carbon hydrogen bond vibrations of aldehydes, alkynes and alkyl groups [6]. The fourth term (1814nm) is attributable to the stretching vibrations of oxygen hydrogen and carbon oxygen bonds in

Table 5.8 Wavelengths contributing to the third season calibration of water activity (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

MLR					
Term	Outliers	$\lambda$ (nm)	$R^2$ (Cal)	SEC	F-value
1	-	1276	0.931	0.0058	1036
2	-	1224	0.943	0.0053	632.5
3	-	1064	0.959	0.0045	577.8
4	-	<b>1814</b>	<b>0.964</b>	<b>0.0043</b>	<b>489.7</b>

cellulose [7], as well as carbon hydrogen bond vibrations of alkyl groups and alkenyl groups with, each with a single hydrogen bonded to it [6]. A plot of the MLR calibration is given above (Figure 5.5). This calibration again demonstrated a similar trend in F-value as the previous seasons, in that a sharp decrease (Table 5.8 Term=1 F-value=1035; Term=4 F-value=490) was seen with the introduction of new terms to the calibration, which may affect the predictions made by this calibration [8]. The final F-value achieved was high (Table 5.8 F-value=490) which, and of comparable size to those seen in previous seasons, which indicates that the calibration is reliable [8]. The coefficient of determination and SEC steadily improved as more terms were introduced to the calibration (Table 5.8 Term=1  $R^2=0.931$ , SEC=0.0058; Term=4  $R^2=0.964$ , SEC=0.0043) and the final values obtained were high as was normally seen with water activity calibrations. These were very good values considering the data were from one seasons' samples.

### 5.2.3.2 PLS calibration

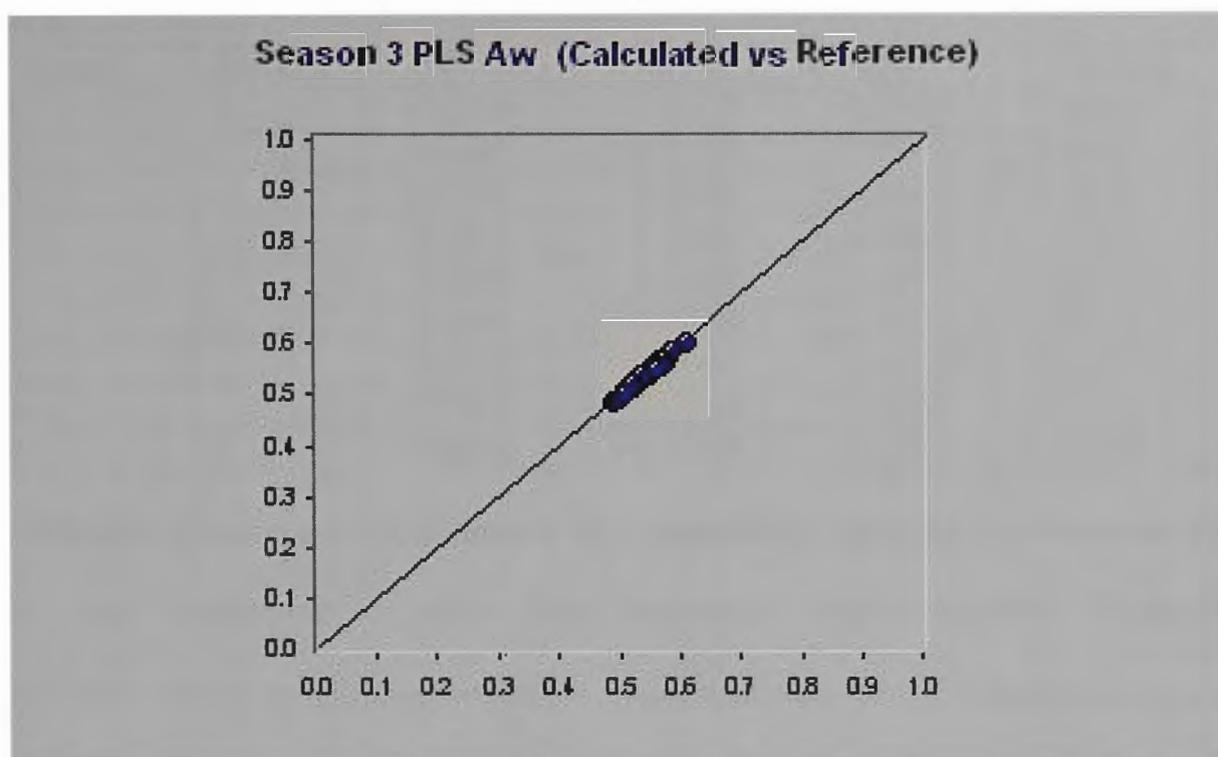


Figure 5.6 A plot of the second seasons' calibration for water activity using PLS

Given above is a plot of the PLS calibration (Figure 5.6). The third seasons' water activity calibrations showed a more typical performance compared to the previous season, PLS (Table 5.9  $R^2=0.972$ ,  $SEC=0.0039$ ) was higher than MLR (Table 5.8  $R^2=0.964$ ,  $SEC=0.0043$ ), though the difference was slight, the difference in the F-value was more marked, with MLR (Table 5.8 F-value 490) likely to be more reliable than PLS (Table 5.9 F-value=265). Again, the PLS calibration showed a general improvement in its F-value (Table 5.9 Factor=1 F-value=49.2; Factor=9 F-value=265), which showed the calibration was not over-fitted by the number of factors used to optimise this calibration [8].

The final result for the PLS calibration was lower than the MLR calibration (Table 5.8 MLR F-value=490; Table 5.9 PLS F-value=265), this may indicate that PLS was more susceptible to spectral outliers than MLR [10] and may have caused this lower F-value for PLS compared to MLR, the final result was a high value and should be able to provide a valid prediction. Error of cross-

Table 5.9 Factors contributing to the third seasons' PLS calibration of water activity (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

PLS					
Factors	Outliers	R <sup>2</sup> (Cal)	SEC	SECV	F-value
1	-	0.390	0.0172	0.0188	49.27
2	-	0.565	0.0146	0.0158	49.37
3	-	0.759	0.0109	0.0124	78.91
4	-	0.914	0.0066	0.0082	196.8
5	-	0.933	0.0059	0.0071	202.7
6	-	0.944	0.0054	0.0067	203.4
7	-	0.957	0.0048	0.0064	226.1
8	-	0.966	0.0042	0.0062	250.2
<b>9</b>	-	<b>0.972</b>	<b>0.0039</b>	<b>0.0061</b>	<b>264.9</b>

validation showed a trend of steady improvement as more factors were added to the calibration (Table 5.9 Factor=1 SECV=0.0188; Factor=9 SECV=0.0061), which demonstrated that predictive model should be robust even when more factors were added to the calibration [9]. As more factors were added to the calibration, both the coefficient of determination and SEC improved markedly (Table 5.9 Factor=1 R<sup>2</sup>=0.390, SEC=0.0172; Factor=9 R<sup>2</sup>=0.972, SEC=0.0039).

## 5.2.4 Combined Seasons Water Activity

Table 5.10 Laboratory parameters summary of combined Aw

# Samples (in calibration)	# Samples in validation	Range	Mean	Standard Deviation
311(249)	62	0.479-0.612	0.543	0.0238

### 5.2.4.1 MLR Calibration

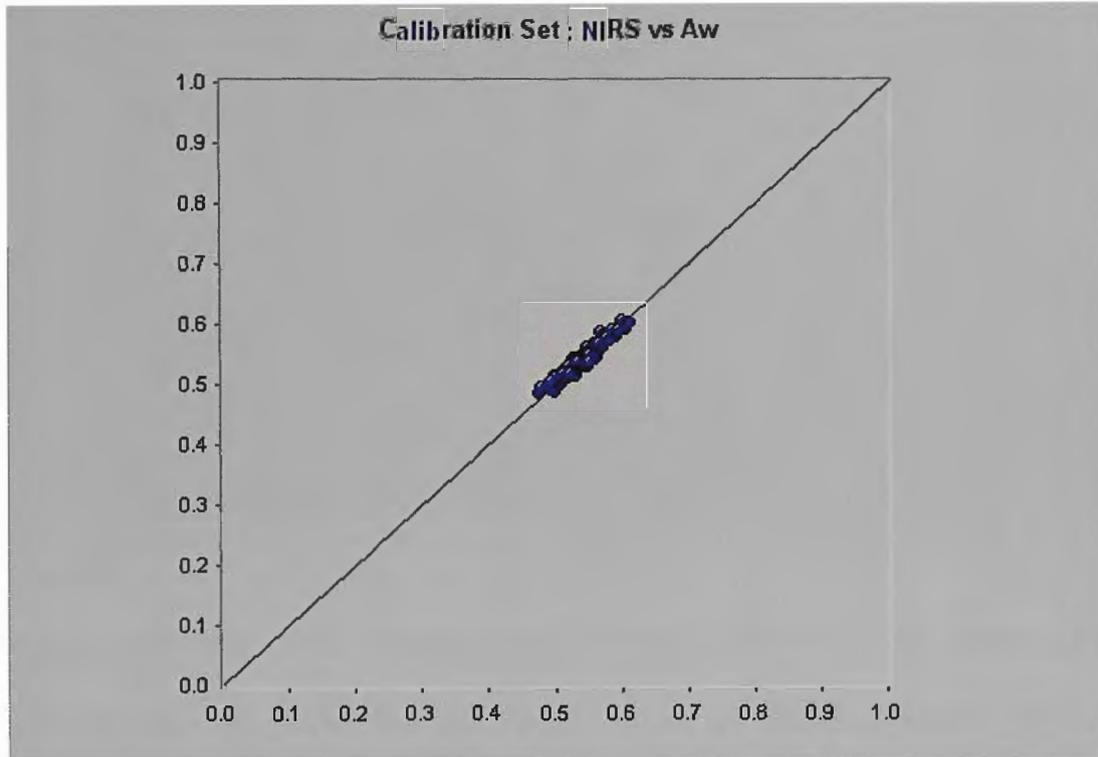


Figure 5.7 A plot of the combined calibration for water activity using MLR

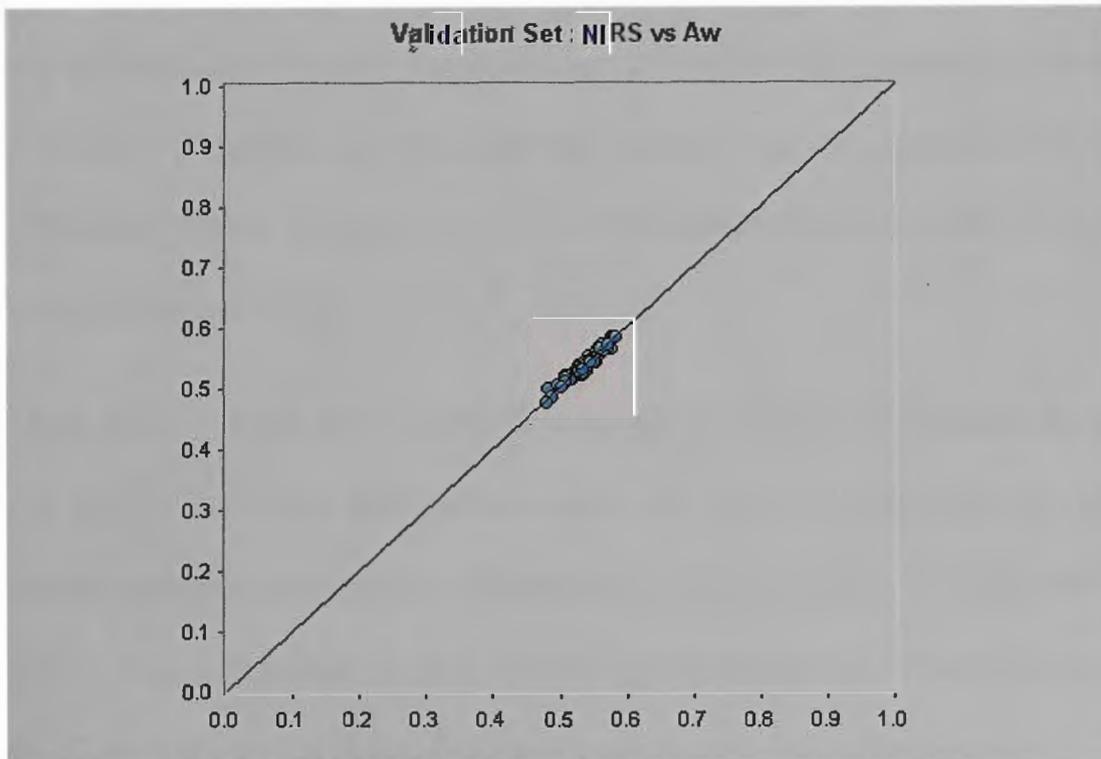


Figure 5.8 A plot of the combined validation set for water activity using MLR

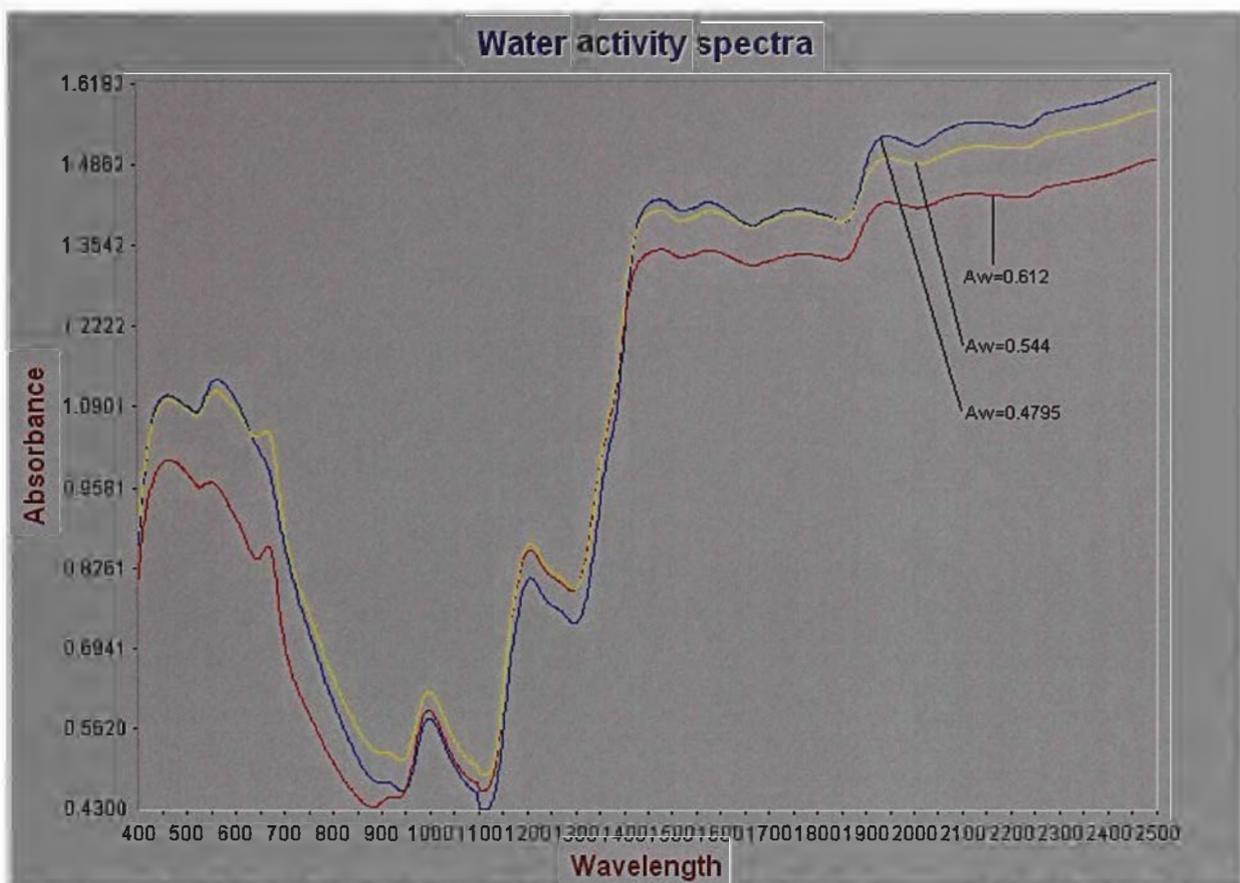


Figure 5.9 Spectra of the minimum (blue), maximum (red) and mean (yellow) samples of water activity

A Diagram showing three spectra from the calibration is given above (Figure 5.9), these spectra show the full region used in the calibration (400nm to 2500nm) and provided are the spectrum of the sample of the maximum water activity value (0.612) in red, the spectrum of the sample closest to the mean value (0.544) in yellow and the spectrum of the minimum sample (0.4795) in blue. Provided below is a close up of the three spectra of the region of the first wavelength (Figure 5.10).

The first term (Table 5.11 1278nm) of the combined calibration of water activity originated from the same region as the first wavelength of the individual seasons calibrations (Tables 5.2, 5.5 and 5.8). The second term (1026nm) was attributed to the stretching vibrations of nitrogen-hydrogen bonds [5] amines and amides, hydroxy groups, and carbon hydrogen

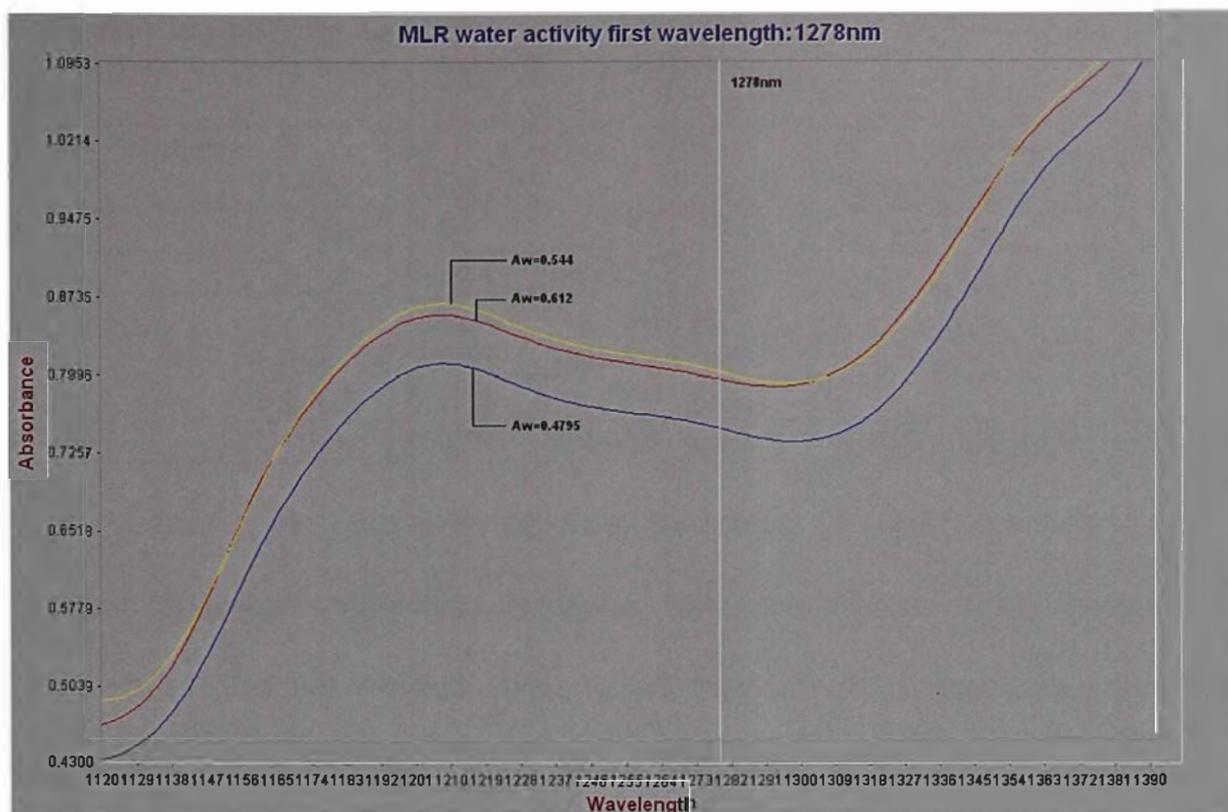


Figure 5.10 Close up of NIR spectra at the first wavelength of the combined MLR calibration.

vibrations in alkyl groups and aldehydes [6]. The third term (1528nm) was linked to the stretching vibrations of oxygen hydrogen bonds [5] of water and hydroxy groups involved in hydrogen bonding, as well as nitrogen hydrogen bond vibrations in amines, primary amides and aqueous ammonia, hydroxy groups, except carboxylic acids [6]. The fourth term (442nm) originated from the visible region of the spectrum [7]. The fifth term (1836nm) was due to the stretching vibrations of oxygen hydrogen and oxygen carbon bonds hydroxy groups within cellulose [7]. The sixth term (1504nm) was traced to first overtone stretching vibrations of nitrogen hydrogen bonds [5] of amines and amides, vibrations within water and hydroxy groups, except carboxylic acids [6]. Plots of the calibration set (Figure 5.7) and the validation set (Figure 5.8) are provided above.

Table 5.11 Wavelengths contributing to the combined seasons' MLR calibration of water activity (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

MLR								
Term	Outliers	$\lambda$ (nm)	$R^2$ (Cal)	SEC	F- value	$R^2$ (Val)	Bias	SEP
1	-	1278	0.89	0.0075	1942	-	-	-
2	-	1026	0.913	0.0068	1197	-	-	-
3	-	1528	0.924	0.0064	921	-	-	-
4	-	442	0.930	0.0061	755	-	-	-
5	-	1836	0.933	0.0060	628	-	-	-
<b>6</b>	-	<b>1504</b>	<b>0.936</b>	<b>0.0059</b>	<b>546</b>	<b>0.965</b>	<b>-0.0008</b>	<b>0.0063</b>

The F-value of the combined seasons calibration of water activity showed a sharper decrease than the individual seasons (Table 5.11 Term=1 F-value=1942; Term=6 F-value=546), which may affect the predictions that this calibration gives; however this was countered by the high final value, indicating a calibration capable of reliable prediction [8]. The fact that the F-value of the combined calibration did not alter to any great extent indicated that the water activity of the sultanas over the three seasons was relatively constant as shown in the range of samples. The calibration was excellent for a non-homogeneous agricultural product, which improved steadily as more terms were added (Table 5.11 Term=1  $R^2=0.89$ , SEC=0.0075; Term=6  $R^2=0.936$ , SEC=0.0056). The validation coefficient of determination achieved by this calibration was 0.965 with an uncorrected SEP of 0.0063 and Bias of -0.0008 (Table 5.11).

### 5.2.4.2 PLS calibration

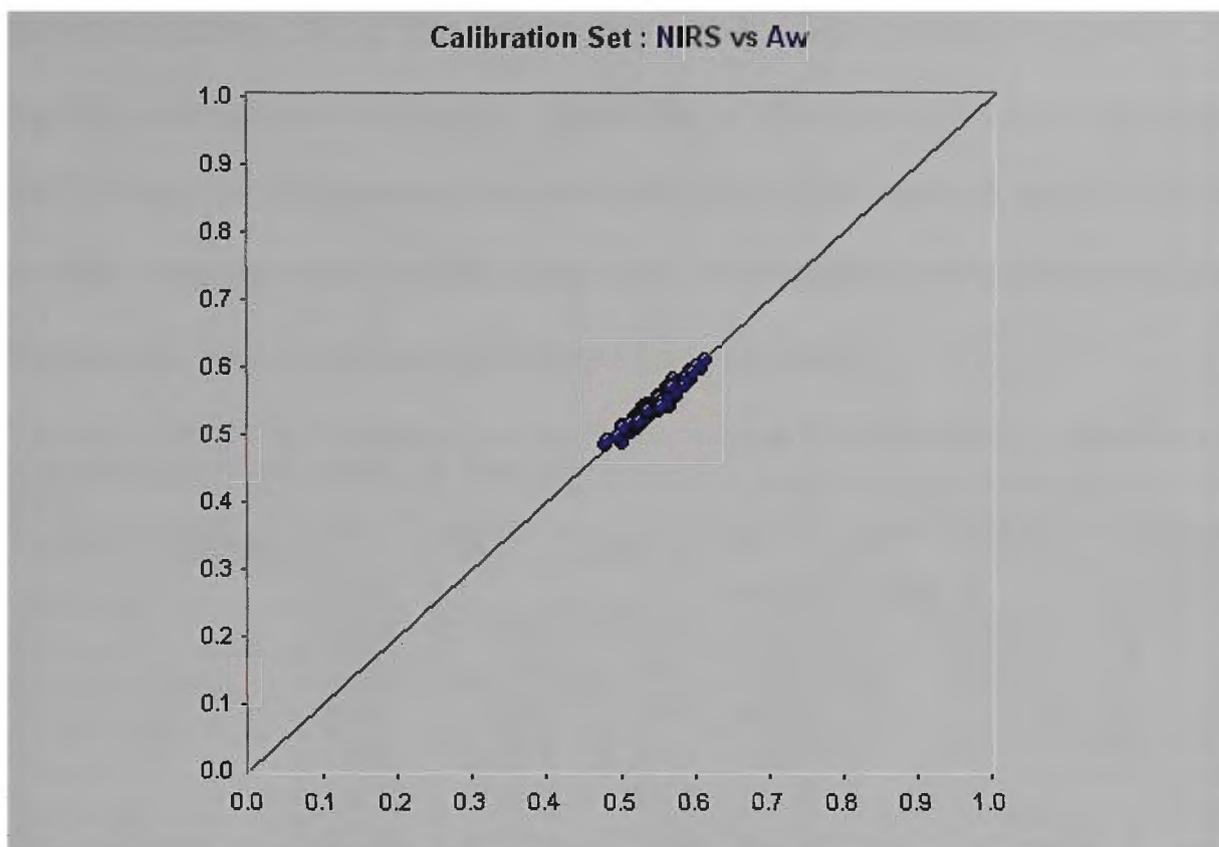


Figure 5.11 A plot of the combined calibration for water activity using PLS

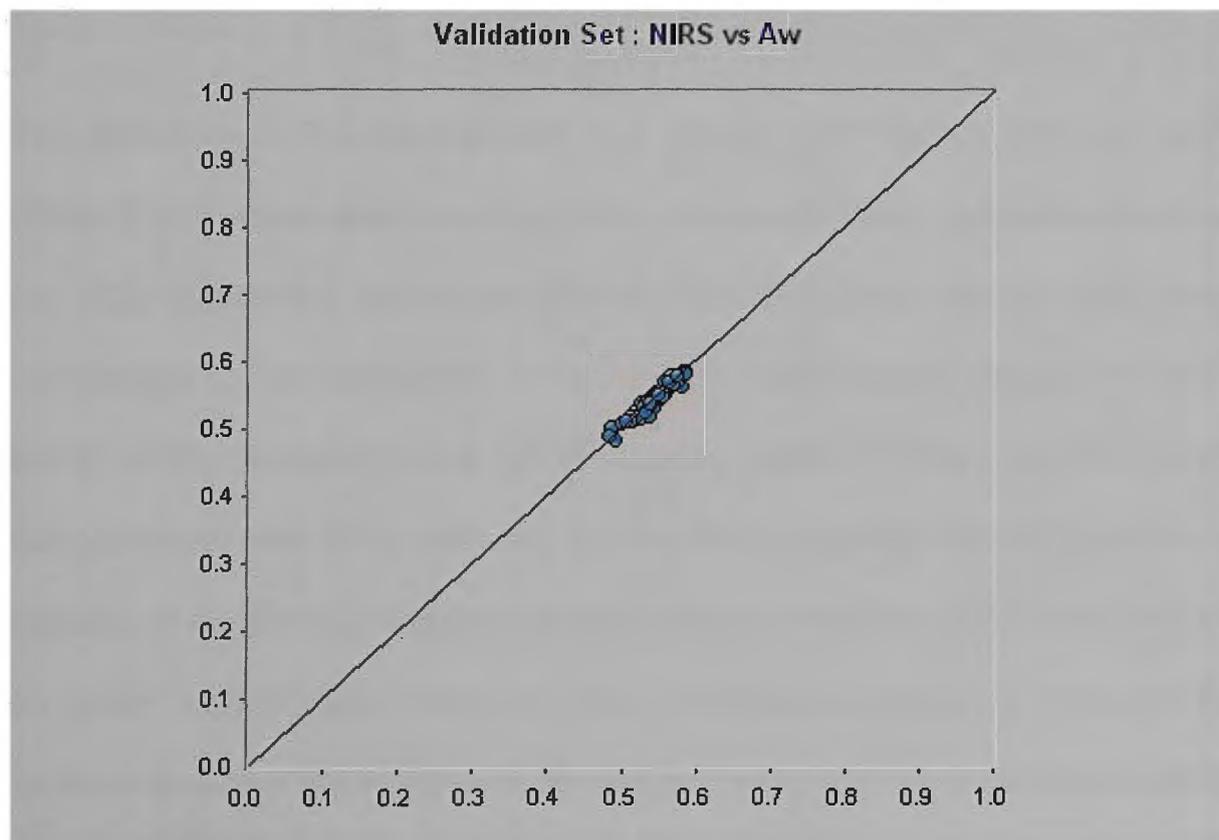


Figure 5.12 A plot of the combined validation set for water activity using PLS

Plots of the calibration set (Figure 5.11) and validation set (Figure 5.12) are provided above. Of all the constituents, water activity was the easiest to achieve a successful calibration regardless of the method used, once again the PLS proved the more robust calibration of the two methods studied. As no outliers were removed from the calibration, the integrity of the calibration and its ability to predict odd samples should be very robust.

Table 5.12 Factors contributing to the combined seasons' PLS calibration of water activity (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

PLS								
Factors	Outliers	R <sup>2</sup> (Cal)	SEC	SECV	F- value	R <sup>2</sup> (Val)	Bias	SEP
1	-	0.584	0.0149	0.0166	323.6	-	-	-
2	-	0.607	0.0145	0.0148	176.9	-	-	-
3	-	0.860	0.0087	0.0092	467.8	-	-	-
4	-	0.890	0.0077	0.0081	460.7	-	-	-
5	-	0.914	0.0068	0.0072	481.7	-	-	-
6	-	0.925	0.0064	0.0069	462.9	-	-	-
7	-	0.936	0.0059	0.0066	470.4	-	-	-
8	-	0.941	0.0057	0.0065	446.9	-	-	-
9	-	0.947	0.0054	0.0063	437.6	-	-	-
<b>10</b>	-	<b>0.949</b>	<b>0.0053</b>	<b>0.0063</b>	<b>410.9</b>	<b>0.957</b>	<b>-0.0013</b>	<b>0.007</b>

The difference in F-value between PLS (Table 5.12 F-value=411) and MLR (Table 5.11 F-value=546) was negligible, considering the improved correlation the PLS combined calibration (Table 5.12 R<sup>2</sup>=0.949 SEC=0.0053) was comparable to the calibrations from the individual seasons (Season 1 Table 5.3 R<sup>2</sup>=0.94; Season 2 Table 5.6 R<sup>2</sup>=0.964; Season 3 Table 5.9 R<sup>2</sup>=0.972), the correlation was less, however, but the final calibration should prove more capable of explaining seasonal variance and is therefore much more robust. As more factors were added, to the combined calibration to account for variance between the spectroscopic and laboratory data both the coefficient of determination and SEC improved (Table 5.12 Factor=1 R<sup>2</sup>=0.584, SEC=0.0149; Factor=10 R<sup>2</sup>=0.949, SEC=0.0053). The PLS calibration

showed a trend in the F-value of a steady increase (Table 5.12 Factor=1 F-value=324; Factor 10 Term=411), which clearly showed that the calibration was not over-fitted by the introduction of the factors used to optimise the calibration [8]. The final F-value (Table 5.12 F-value=411) was much larger than those of the individual seasons PLS calibrations (Table 5.3 Season 1 F-value=253; Table 5.6 Season 2 F-value=239; Table 5.9 Season 3 F-value=265), which would indicate the combination of all three seasons results has improved the reliability of the calibration. SECV steadily improved as more factors were added to the calibration (Table 5.12 Factor=1 SECV=0.0166; Factor=10 SECV=0.0063) indicating the prediction made by the calibration was reliable [9], which was supported by the coefficient of determination of the validation set (Table 5.12 Validation  $R^2=0.957$ ). The coefficient of determination of the validation set achieved was 0.957 with an uncorrected SEP of 0.007 and a Bias of -0.0013 (Table 5.12).

### **5.2.5 Comparison to cited literature**

Work by Huxsoll [11,12] developed a series of NIRS calibrations to assess water activity; in raisins and compared them to values obtained by a cooled mirror dew point apparatus. MPLS was used to develop the model and it was validated with cross validation. The combined seasons calibration developed using both PLS and MLR has yielded results comparable to those obtained by Huxsoll, even though the comparatively limited range of the samples. The correlation coefficients obtained (Table 5.11 MLR  $R^2=0.936$ ; Table 5.12 PLS  $R^2=0.949$ ) were less than those of the individual seasons (Season 1 MLR  $R^2=0.939$ , PLS  $R^2=0.94$ ; Season 2 MLR  $R^2=0.971$ , PLS  $R^2=0.964$ ; Season 3

MLR  $R^2=0.964$ , PLS  $R^2=0.972$ ) as well as that obtained by the raisin calibration (Huxsoll  $R^2=0.979$ ), however the cross validation error achieved was lower (Combined PLS SECV=0.0063; Huxsoll SECV=0.014), which coupled with the seasonal variety included in the calibration indicates it was a more valid model, particularly as the quality of sultanas varies from one season to another the fact that sultanas are a non-homogeneous agricultural product [11,12].

There are few reports in the literature of projects using NIRS to predict Dew Point water activity; some projects assessing related constituents, for example moisture content, make a useful comparison. Work by Flinn *et al* [13] developed calibrations to assess constituents with NIRS of whole and ground pulse samples. Moisture content of ground chickpeas and field peas achieved coefficients of determination of 0.90 and 0.97 respectively and whole chickpea and field pea coefficients of determination of 0.51 and 0.95 respectively. This compared to the results obtained with processed sultanas have given coefficients of determination comparable with the work by Flinn *et al* [13].

Blakeney *et al* [14] developed a calibration to assess moisture contents in the nine commercially available rice varieties grown in the industry using near infrared transmission spectroscopy achieving a coefficient of determination of 0.993 for the initial calibration. Due to the highly homogeneous nature of rice compared to processed sultanas this fact explains the lower coefficient of determination obtained (Table 5.11 MLR  $R^2=0.936$ ; Table 5.12 PLS  $R^2=0.949$ ) [14].

## 5.3 Kjeldahl protein

### 5.3.1 Season one Kjeldahl Protein

Table 5.13 Laboratory parameters summary of season one Kjeldahl protein

# Samples (in calibration)	# Samples in validation	Range	Mean	Standard Deviation
35(28)	7	1.31-3.5	2.29	0.550

#### 5.3.1.1 MLR Calibration

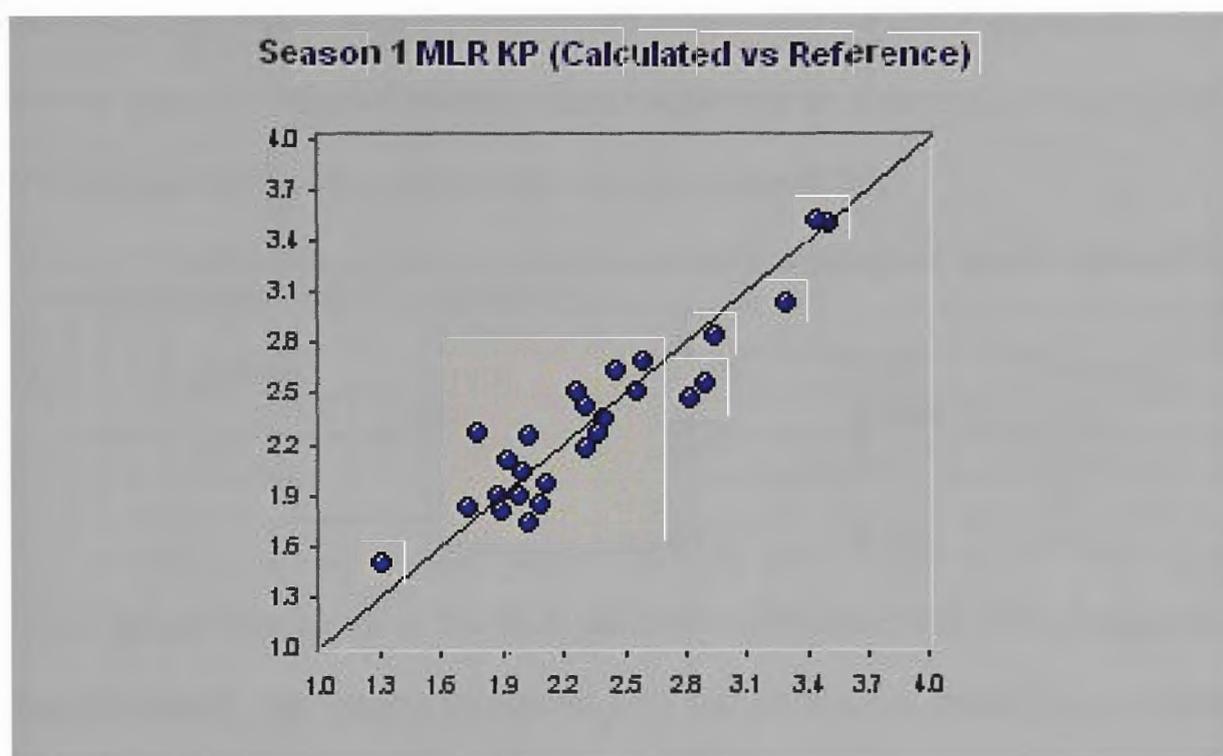


Figure 5.13 A plot of the first seasons' calibration for Kjeldahl protein using MLR

The initial summation term (Table 5.14 468nm) of this calibration originated from the visible region of the spectra [7], and considering that nitrogen containing compounds such as arginine act as substrates for browning processes [1,2], this was quite plausible. The second term (2126nm) was due to both nitrogen-hydrogen single bond and carbonyl double bond stretching vibrations, often due to the presence of amino acids and secondary amines, also *cis* and *trans* alkene and aryl carbon hydrogen vibrations, hydroxy groups, except aryl hydroxides, carboxylates, esters, anhydrides, ketones and

water all absorb at this wavelength [6]. The calibration third term (1414nm) was associated with secondary amides, as well as stretching and deforming vibrations of carbon-hydrogen bonds of alkyl and alkenyl groups except terminal alkenyl groups, hydroxy group containing compounds and water [6]. The asymmetric stretching vibrations of nitrogen-hydrogen single bonds of amide groups are associated with the fourth summation term (1986nm) of this calibration [6], such features were often associated with protein also terminal alkenyl groups, primary amines, aqueous ammonia, hydroxy groups (except primary alcohols) and water absorb within this region [7].

Table 5.14 Wavelengths contributing to the first seasons' calibration of Kjeldahl protein (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

MLR					
Term	Outliers	$\lambda$ (nm)	$R^2$ (Cal)	SEC	F-value
1	-	468	0.508	0.389	24.8
2	-	2126	0.697	0.315	25.6
3	-	1414	0.806	0.255	30.1
4	-	1986	0.861	0.221	32.6

Given above was a plot of the MLR calibration (Figure 5.13). The F-value for this calibration has slowly increased with the addition of more terms (Table 5.14 Term=1 F-value=24.8; Term=4 F-value=32.6) indicating that the calibrations reliability did not deteriorate as more terms were added though the final value was small which indicated that the calibration would not provide valid predictions [8], however this was a small sample set that represents a preliminary calibration. The inclusion of more terms to the calibration resulted in a sharp improvement of the coefficient of determination and SEC (Table 5.14 Term=1  $R^2$ =24.8, SEC=0.389; Term=4  $R^2$ = 0.861, SEC=0.221).

### 5.3.1.2 PLS calibration

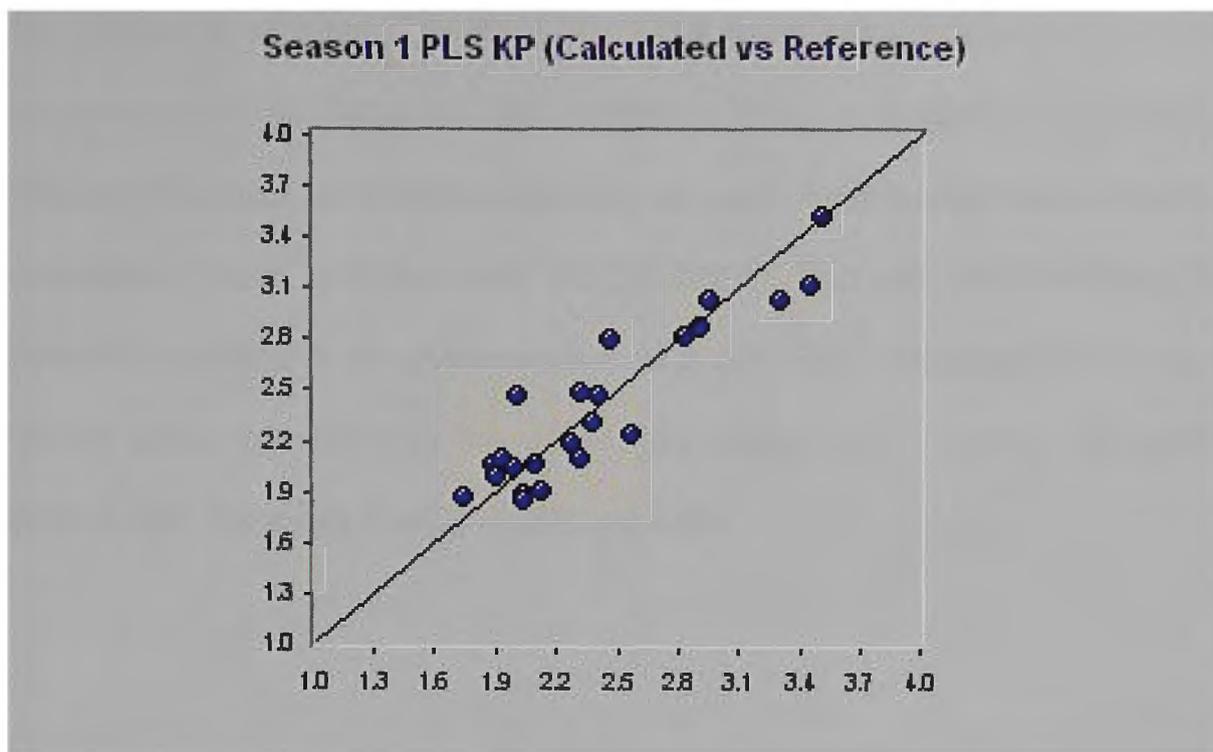


Figure 5.14 A plot of the first seasons' calibration for Kjeldahl protein using PLS

A plot of the calibration is provided above (Figure 5.14). The calibrations developed from the first seasons' Kjeldahl results have not followed previous cases, and the MLR calibration had a slightly higher  $R^2$  value so indicated improved performance to the PLS calibration (Table 5.15 PLS  $R^2=0.848$ ; Table 5.14 MLR  $R^2=0.861$ ), the F-values of both calibrations are also not behaving typically as the MLR F-value (Table 5.14 F-value=32.6) is lower than PLS (Table 5.15 F-value=35.3).

Table 5.15 Factors contributing to the first seasons' PLS calibration of Kjeldahl protein (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

PLS					
Factors	Outliers	$R^2$ (Cal)	SEC	SECV	F-value
1	3	0.212	0.467	0.501	5.66
2	3	0.788	0.248	0.386	37.20
3	3	<b>0.848</b>	<b>0.216</b>	<b>0.326</b>	<b>35.29</b>

Such unusual results were attributed to the small sample set. The F-value of the PLS calibration clearly indicates it is preliminary data, the value is low but the value itself increases with the number of factors that were introduced

(Table 5.15 Factor=1 F-value=5.66; Factor=3 F-value=35.29), this implied that the calibration was not over-fitted [8]. The trend seen in the error of cross validations also indicates that the calibration is not over-fitted; the prediction assessed in cross validation improves as more factors were added to the calibration (Table 5.15 Factor=1 SECV=0.501; Factor=3 SECV=0.326) [9]. Both the coefficient of determination and the SEC improved when more factors were included into the calibration (Table 5.15 Factor=1  $R^2=0.212$ , SEC=0.467; Factor=3  $R^2=0.848$ , SEC=0.216).

### 5.3.2 Season two Kjeldahl protein

Table 5.16 Laboratory parameters summary of season two Kjeldahl Nitrogen

# Samples (in calibration)	# Samples in validation	Range	Mean	Standard Deviation
95 (76)	19	1.82-3.37	2.46	0.336

#### 5.3.2.1 MLR calibration

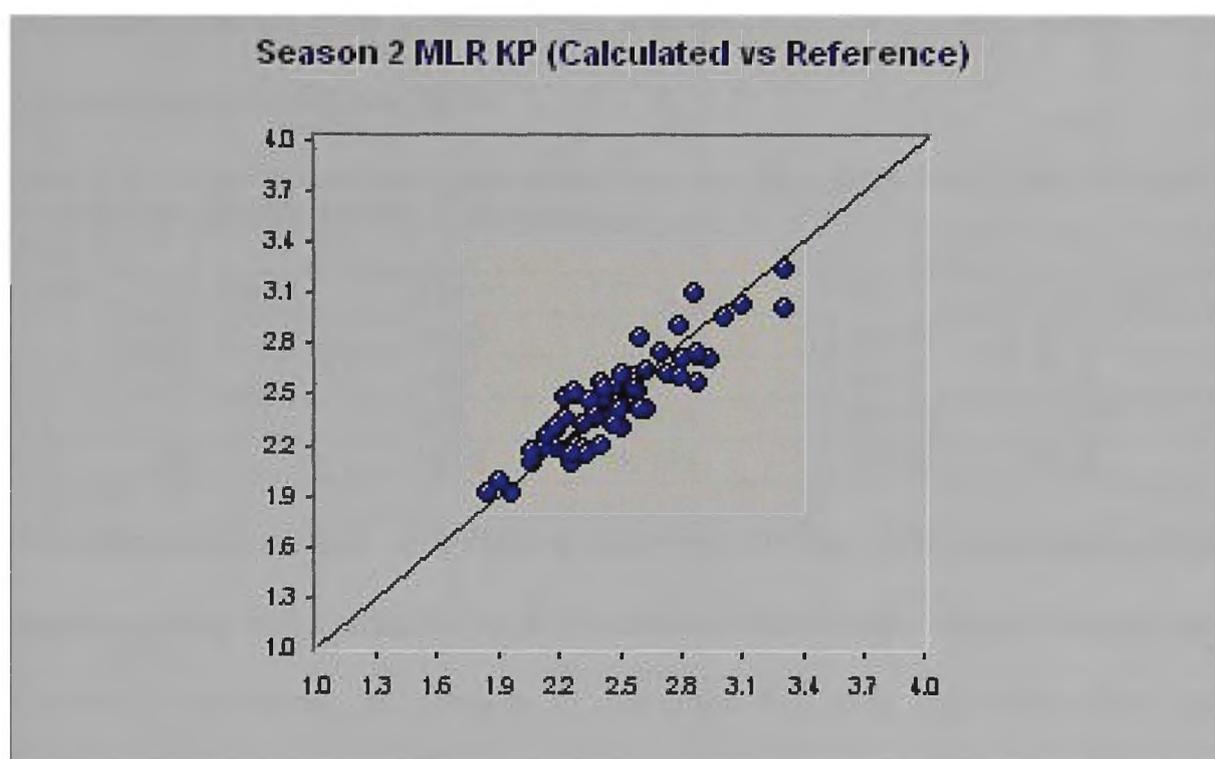


Figure 5.15 A plot of the second seasons' calibration for Kjeldahl protein using MLR

The first term of the calibration (Table 5.17 1226nm) of Kjeldahl protein of the second season was linked to the second overtone stretching vibrations of carbon-hydrogen bonds [5], typically caused by alkyl groups and alkenyl groups with one hydrogen bonded to it [7]. The next two terms (488nm and 658nm) were associated with the visible region of the spectra [6]. The last term (1412nm) can be attributed to secondary amides as well as the first overtone stretching vibrations of oxygen-hydrogen bonds of hydroxy groups [5], as well as alkyl, aldehyde, alkenyl carbon hydrogen bonds absorb at this wavelength [6,7]. Provided above is a plot of the MLR calibration (Figure 5.15). The trend seen in F-value for this seasons calibration showed a steady

increase as more terms are added to the calibration (Table 5.17 Term=1 F-value=34; Term=4 F-value=58.8) [8]. The calibration was less strongly correlating than the first season calibration (Table 5.14 Season 1  $R^2=0.86$ , SEC=0.22; Table 5.17 Season 2  $R^2=0.81$ , SEC=0.14), despite it being larger and containing a better distribution of samples along the full range of the calibration, the error of calibration of the second season was lower which indicates that this was the case.

Table 5.17 Terms contributing to the second seasons' MLR calibration of Kjeldahl protein (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

MLR					
Term	Outliers	$\lambda$ (nm)	$R^2$ (Cal)	SEC	F-value
1	8	1226	0.365	0.249	34.0
2	"	488	0.591	0.202	41.8
3	"	658	0.698	0.175	43.9
<b>4</b>	<b>8</b>	<b>1412</b>	<b>0.808</b>	<b>0.141</b>	<b>58.8</b>

This disparity between calibrations was due to the calibration of the first season giving an artificially high correlation due to the small sample set (Season 1: Samples= 35. Season 2; Samples= 95), also the first season was more easily affected by outliers due to its small size, which results in a higher error of calibration. The coefficient of determination and SEC obtained by this calibration improved as more terms were added to the calibration (Table 5.17 Term=1  $R^2=0.365$ , SEC=0.249; Term=4  $R^2=0.808$ , SEC=0.141).

### 5.3.2.2 PLS Calibration

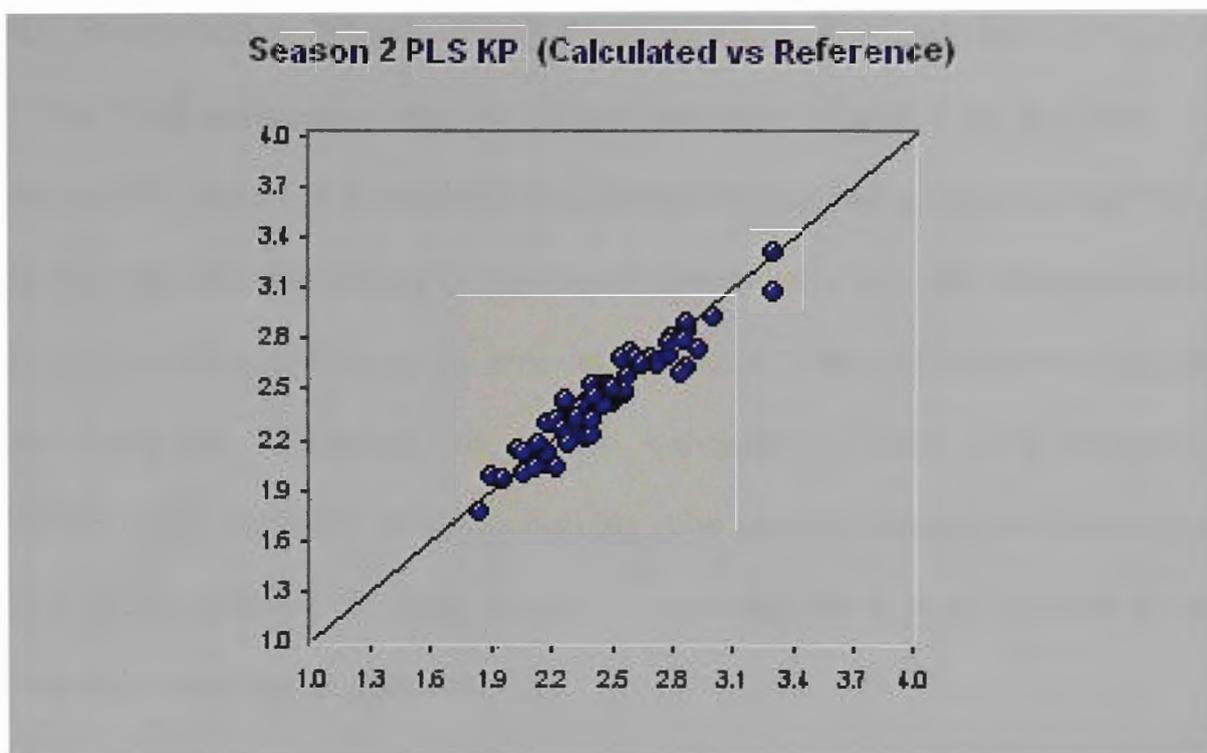


Figure 5.16 A plot of the second seasons' calibration for Kjeldahl protein using PLS

Provided above is a plot of the PLS calibration (Figure 5.16). The second seasons' calibrations of Kjeldahl protein showed more typical behaviour compared to the first season, The PLS calibration achieved a marked improvement in performance compared to MLR, as can be seen with the corresponding correlation and error of calibration (Table 5.17 MLR  $R^2=0.81$ ,  $SEC=0.14$ ; Table 5.18 PLS  $R^2=0.90$ ,  $SEC=0.10$ ).

Table 5.18 Factors contributing to the second seasons' PLS calibration of Kjeldahl protein (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

PLS					
Factors	Outliers	$R^2$ (Cal)	SEC	SECV	F-value
1	8	0.059	0.298	0.302	3.75
2	"	0.321	0.255	0.278	13.94
3	"	0.541	0.211	0.252	22.81
4	"	0.656	0.185	0.239	27.12
5	"	0.764	0.154	0.211	36.29
6	"	0.807	0.140	0.205	38.37
7	"	0.860	0.121	0.194	47.52
<b>8</b>	<b>8</b>	<b>0.898</b>	<b>0.104</b>	<b>0.194</b>	<b>58.12</b>

The difference in F-values of the seasons' calibrations was slight (Table 5.17 MLR F-value=58.8; Table 5.18 PLS F-value= 58.2). The trend seen in F-value of the PLS calibration was of steady increase (Table 5.18 Factor=1 F-value=3.75; Factor=8 F-value=58.12) indicating that there was no over-fitting [8], this was also supported by the trend seen in error of cross validation as it showed a steady decline in the amount of variation from the reference data by the prediction assessed by cross validation (Table 5.18 Factor=1 SECV=0.302; Factor=8 SECV=0.194) [9]. The second season calibration was more representative of crude protein in sultanas, as it was a much larger calibration than the first season.

### 5.3.3 Season three Kjeldahl protein

Table 5.19 Laboratory parameters summary of season three Kjeldahl Nitrogen

# Samples (in calibration)	# Samples in validation	Range	Mean	Standard Deviation
156(125)	31	1.48-3.20	2.36	0.317

#### 5.3.3.1 MLR Calibration

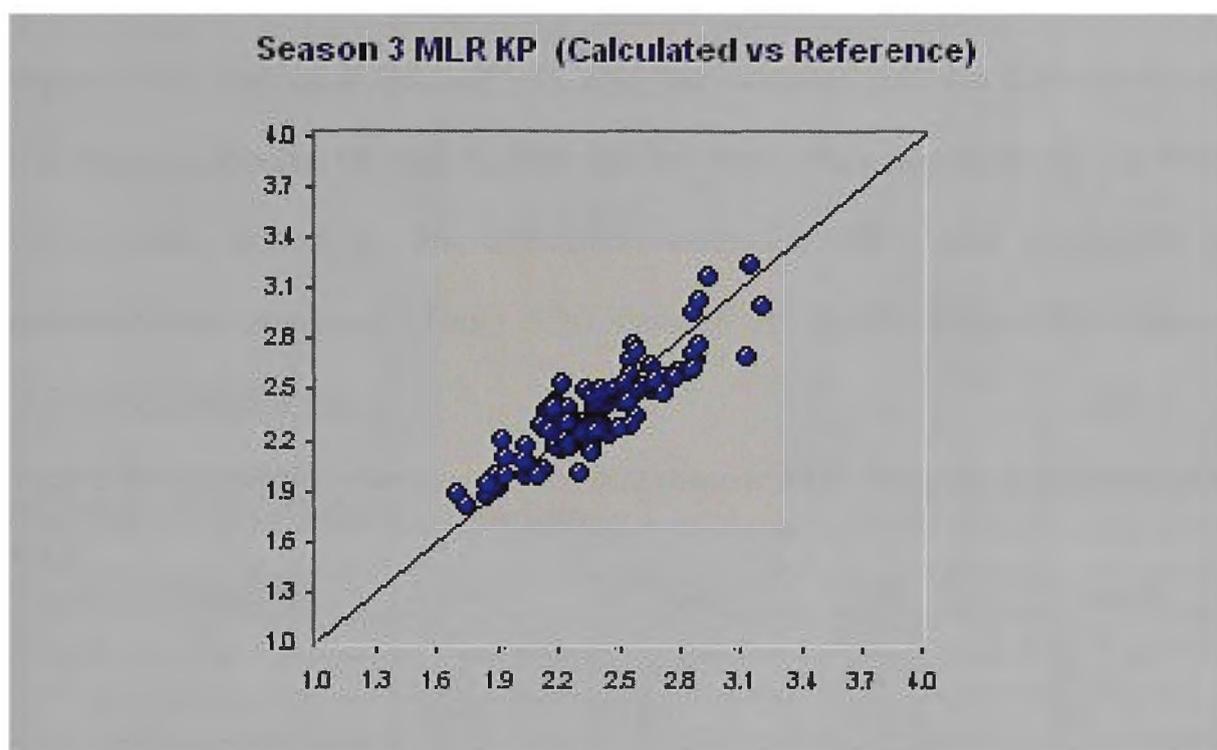


Figure 5.17 A plot of the third seasons' calibration for Kjeldahl protein using MLR

The first two terms (Table 5.20 584nm and 620nm) of the Kjeldahl protein calibration of the third season came from the visible region of the spectrum [7]. The third term (1254nm) came from near the first overtone carbon hydrogen bond combinations band [5]. The fourth term (1222nm) was due to the second overtone stretching vibrations of carbon hydrogen bonds [5] of alkyl and alkenyl groups with one hydrogen bonded to them [6]. The fifth and sixth terms (1180nm and 1200nm) of the calibration are attributable to second overtone stretching vibrations of carbon hydrogen bonds [5]. At 1180nm, alkyl and alkenyl, except terminal alkenyl groups absorb, while at 1200nm,

alkyl groups, trans alkenes and alkenes with one carbon to hydrogen bond absorb [6]. A plot of the MLR calibration is provided above (Figure 5.17). The F-value of the second seasons' calibration was less when compared to the previous season (Season 2 Table 5.17 F-value=58.8; Season 3 Table 5.20 F-value=51.6), it also showed steady improvement after an initial sharp decline as more terms were added to the calibration (Table 5.20 Term=1 F-value=91.5; Term=6 F-value=51.6), and this indicated that the calibration was not notably affected by optimisation as the final value was high [8]. As more terms were added to the calibration both the SEC and coefficient of determination improved (Table 5.20 Term=6  $R^2=0.765$ , SEC=0.15; Term=6  $R^2=0.765$ , SEC=0.15).

Table 5.20 Wavelengths contributing to the third seasons' MLR calibration of Kjeldahl protein (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

MLR					
Term	Outliers	$\lambda$ (nm)	$R^2$ (Cal)	SEC	F-value
1	13	584	0.478	0.218	91.5
2	"	620	0.545	0.204	59.3
3	"	1254	0.60	0.193	49
4	"	1222	0.645	0.1825	44
5	"	1188	0.697	0.1694	44.2
<b>6</b>	<b>13</b>	<b>1200</b>	<b>0.765</b>	<b>0.150</b>	<b>51.6</b>

The calibration of the third season achieved a lower coefficient of determination than the first two seasons' MLR calibrations (Season 1 Table 5.14  $R^2=0.86$ , SEC=0.22; Season 2 Table 5.17  $R^2=0.81$ , SEC=0.14; Season 3 Table 5.20  $R^2=0.765$ , SEC=0.15) which was due to the preliminary nature of this calibration, and a comparable error of calibration to the second season, which indicated it was a more representative calibration than the first season as both the second and final season were less susceptible to outliers due to their size [15].

### 5.3.3.2 PLS calibration

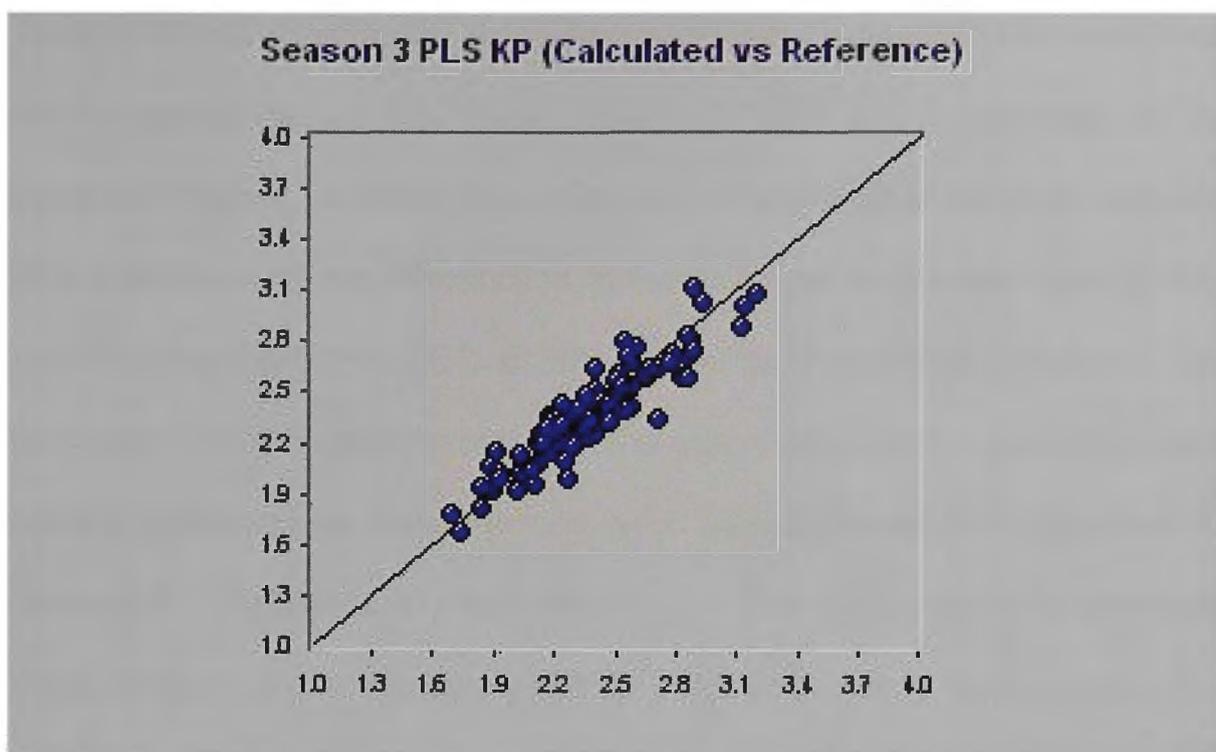


Figure 5.18 A plot of the third seasons calibration for Kjeldahl protein using PLS

A plot of the PLS calibration is provided above (Figure 5.18). The calibrations developed from the Kjeldahl nitrogen results from the third season of the project showed a higher coefficient of determination developed by PLS compared to MLR, and a correspondingly lower error of calibration (Table 5.20 MLR  $R^2=0.765$ ,  $SEC=0.15$ ; Table 5.21 PLS  $R^2=0.84$ ,  $SEC=0.129$ ).

Table 5.21 Factors contributing to the third seasons PLS calibration of Kjeldahl protein (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

PLS					
Factors	Outliers	$R^2$ (Cal)	SEC	SECV	F-value
1	13	0.209	0.268	0.273	26.37
2	"	0.351	0.244	0.256	26.82
3	"	0.491	0.217	0.242	31.58
4	"	0.516	0.213	0.240	25.88
5	"	0.553	0.206	0.243	23.78
6	"	0.600	0.196	0.248	23.76
7	"	0.671	0.178	0.249	27.42
8	"	0.703	0.170	0.247	27.49
9	"	0.742	0.160	0.239	29.40
10	"	0.770	0.152	0.240	30.41
11	"	0.798	0.143	0.242	32.42
12	"	0.824	0.134	0.237	34.79
<b>13</b>	<b>13</b>	<b>0.840</b>	<b>0.129</b>	<b>0.235</b>	<b>35.54</b>

The F-value of the PLS calibration was low compared to MLR of this season (Table 5.20 MLR F-value=51.6; Table 5.21 PLS F-value=35.5) and was lower than its second season equivalent (Season 2 Table 5.18 F-value=58.12) but this was, however, a preliminary calibration. Coefficient of determination and SEC improved with the introduction of more factors to the calibration (Table 5.21 Factor=1  $R^2=0.209$ , SEC=0.268; Factor=13  $R^2=0.84$ , SEC=0.129). The trend seen in the F-value shows that the calibration was not over-fitted as it steadily increased as more factors were added (Table 5.21 Factor=1 F-value=26.37; Factor=13 F-value=35.54) [8]. The trend seen in the error of cross validation showed a steady decrease as more factors were added to the calibration (Table 5.21 Factor=1 SECV=0.273; Factor=13 SECV=0.235), indicating the prediction made by the calibration improved as more factors were added [9].

### 5.3.4 Combined seasons Kjeldahl protein

Table 5.22 Laboratory parameters summary of combined Kjeldahl protein

# Samples (in calibration)	# Samples in validation	Range	Mean	Standard Deviation
186(149)	37	1.31-3.5	2.39	0.428

#### 5.3.4.1 MLR calibration

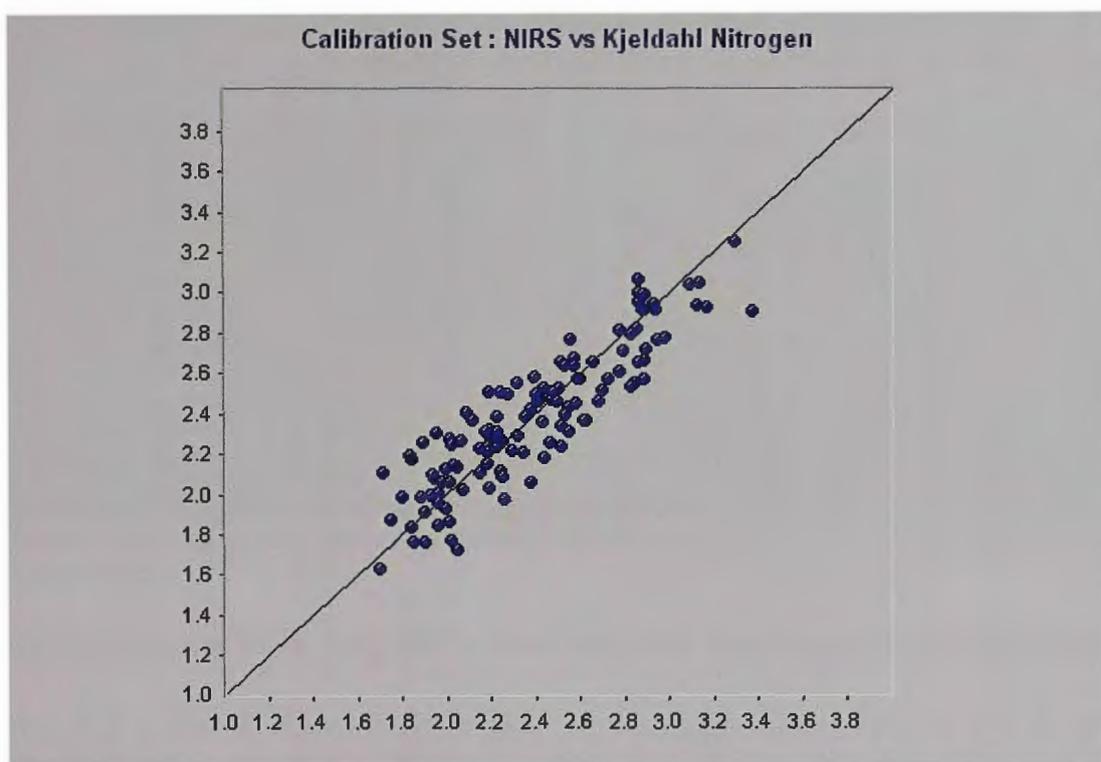


Figure 5.19 A plot of the combined calibration of Kjeldahl protein using MLR

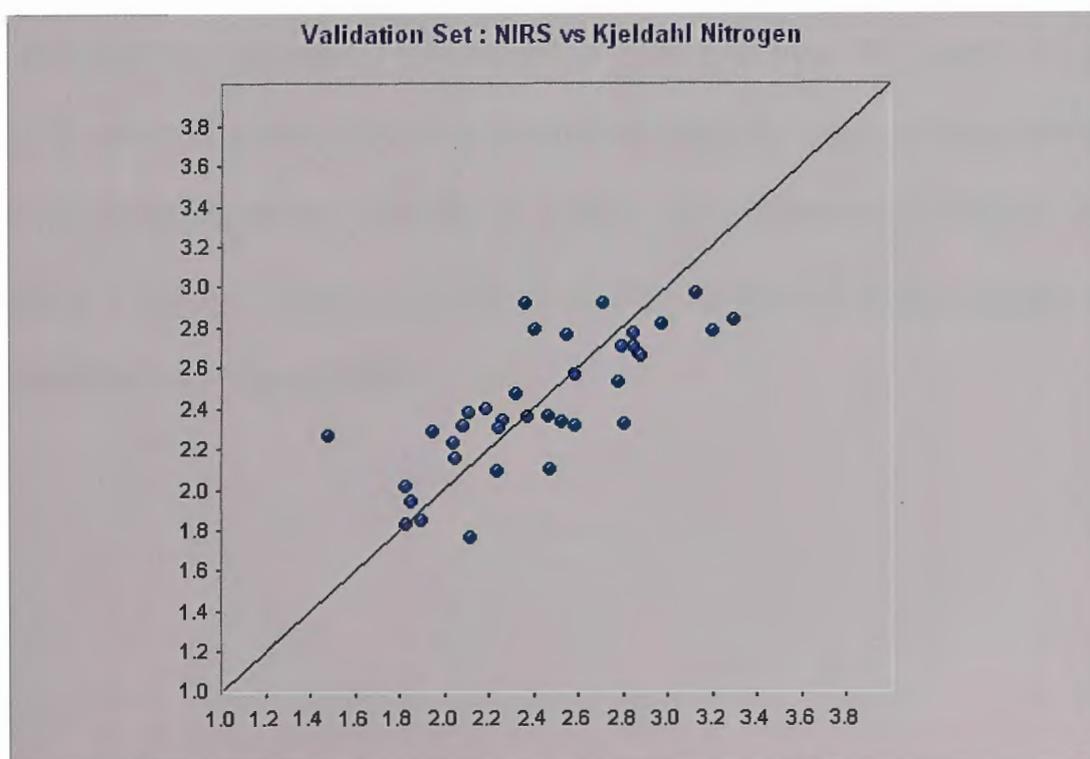


Figure 5.20 A plot of the combined validation set for protein using MLR

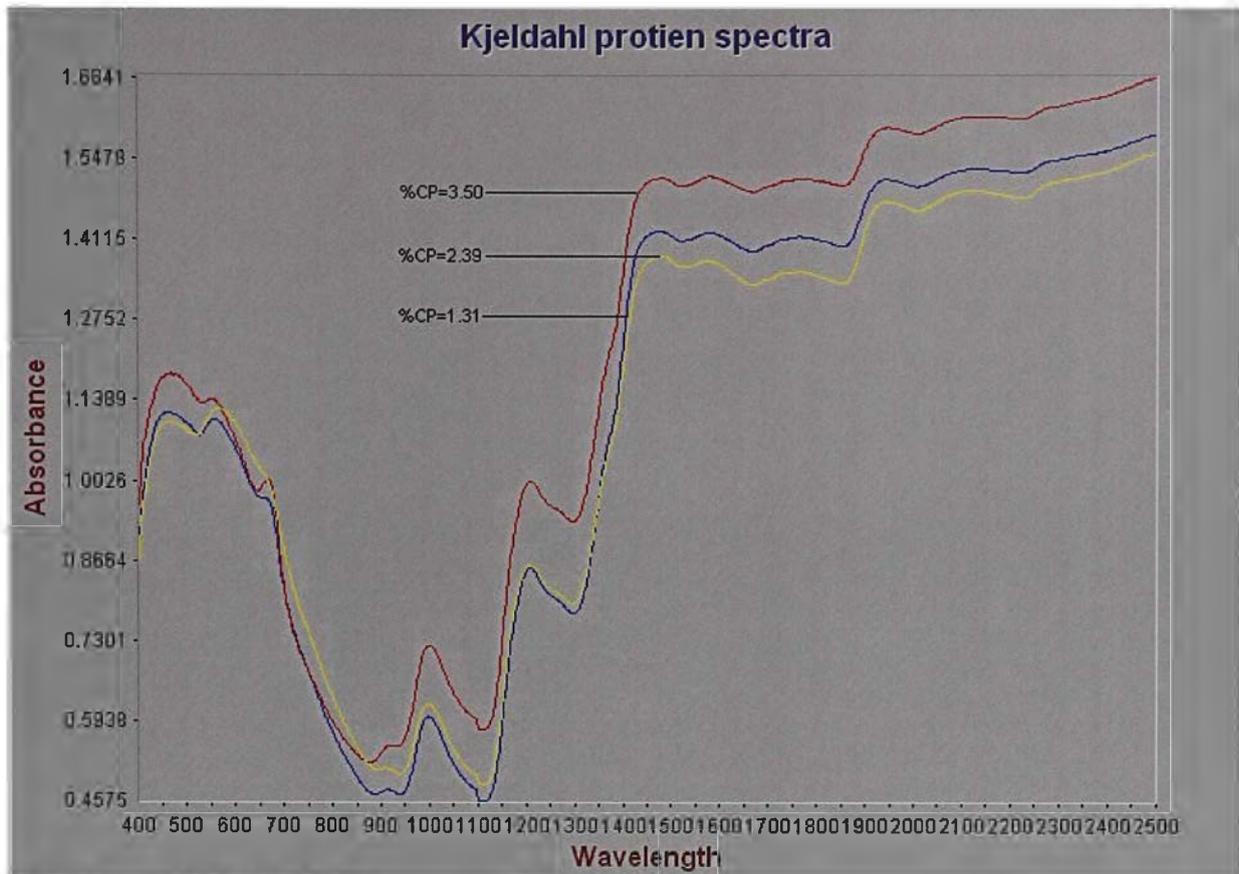


Figure 5.21 Spectra of the minimum (blue), maximum (red) and mean (yellow) samples of Kjeldahl protein

Presented above is a diagram showing three spectra of the calibration set (Figure 5.21) these spectra show the whole region covered by the spectrum used in the calibration (400 to 2500 nm) and are the spectrum of the sample that yielded the minimum Kjeldahl value (1.31) in blue the maximum value (3.50) in red and a spectrum of a laboratory value as close as possible to the mean value (2.39) of the sample set used in this calibration in yellow. Below is given a close up of those spectra at the region of the first wavelength of the MLR calibration (Figure 5.22).

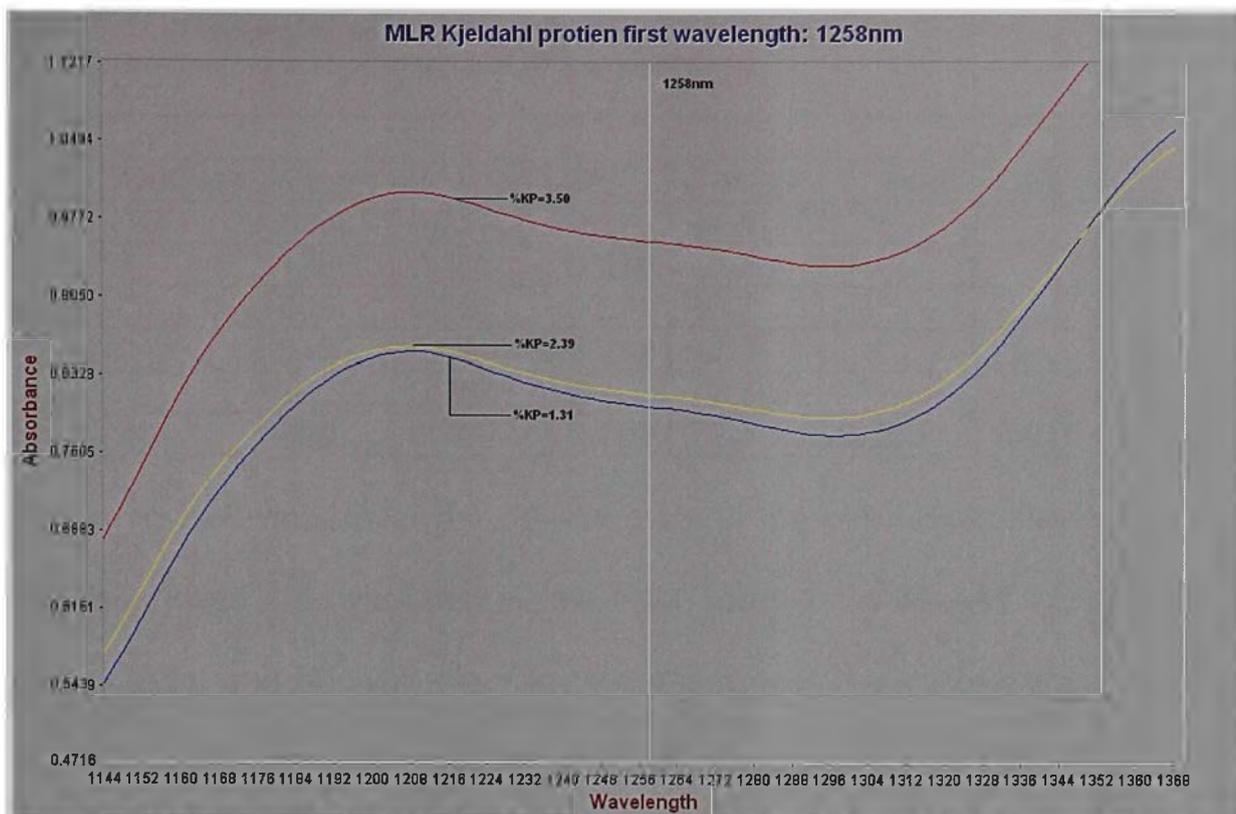


Figure 5.22 Close up of NIR spectra at the first wavelength of the combined MLR calibration

The first two terms (Table 5.23 1256nm and 1286nm) of the combined calibration of Kjeldahl Nitrogen originated from the first overtone carbon hydrogen combinations region [5]. The third term (808nm) was attributable to the stretching and deformation vibrations of nitrogen hydrogen bonds and the stretching vibrations of carbon nitrogen bonds of secondary amines [6]. The fourth term (1184nm) was linked to the second overtone stretching vibrations of carbon hydrogen bonds [5] of alkyl and alkenyl groups (except terminal alkenyl groups) [6]. The fifth term (1220nm) was attributed to the stretching vibrations of carbon hydrogen bonds of alkyl groups and alkenyl groups with a single carbon hydrogen bond [6]. The sixth term (800nm) was due to the third overtone stretching vibrations of nitrogen hydrogen bonds [5] of primary amines [6]. Plots of the calibration set (Figure 5.19) and validation set (Figure 5.20) of the MLR calibration are provided above.

Table 5.23 Wavelengths contributing to the combined seasons MLR calibration of Kjeldahl protein (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

MLR								
Term	Outliers	$\lambda$ (nm)	R <sup>2</sup> (Cal)	SEC	F-value	R <sup>2</sup> (Val)	Bias	SEP
1	15	1258	0.274	0.33	44.5	-	-	-
2	"	1286	0.408	0.299	40.4	-	-	-
3	"	808	0.507	0.274	39.7	-	-	-
4	"	1184	0.596	0.249	42.5	-	-	-
5	"	1220	0.761	0.193	72.6	-	-	-
<b>6</b>	<b>15</b>	<b>800</b>	<b>0.793</b>	<b>0.180</b>	<b>72.1</b>	<b>0.759</b>	<b>0.016</b>	<b>0.279</b>

The trend in F-value over the development of the calibration shows that the calibration was not over-fitted (Table 5.23 Term=1 F-value=44.5; Term=6 F-value=72.1), and showed clear improvement over the calibrations from the individual seasons (Table 5.23 Combined seasons F-value=72.1; Table 5.14 Season 1 F-value=32.6; Table 5.17 Season 2 F-value=58.8; Table 5.20 Season 3 F-value=51.6). The calibration should thus provide reliable predictions [8]. As more terms were added to the calibration both the coefficient of determination and SEC improved sharply (Table 5.23 Term=1 R<sup>2</sup>=0.274, SEC=0.33; Term=6 R<sup>2</sup>=0.793, SEC=0.180), and successfully predicted the validation set (Table 5.23 Validation R<sup>2</sup>=0.759). The validation set achieved an uncorrected SEP of 0.279 and Bias of 0.016 (Table 5.23).

### 5.3.4.2 PLS Calibration

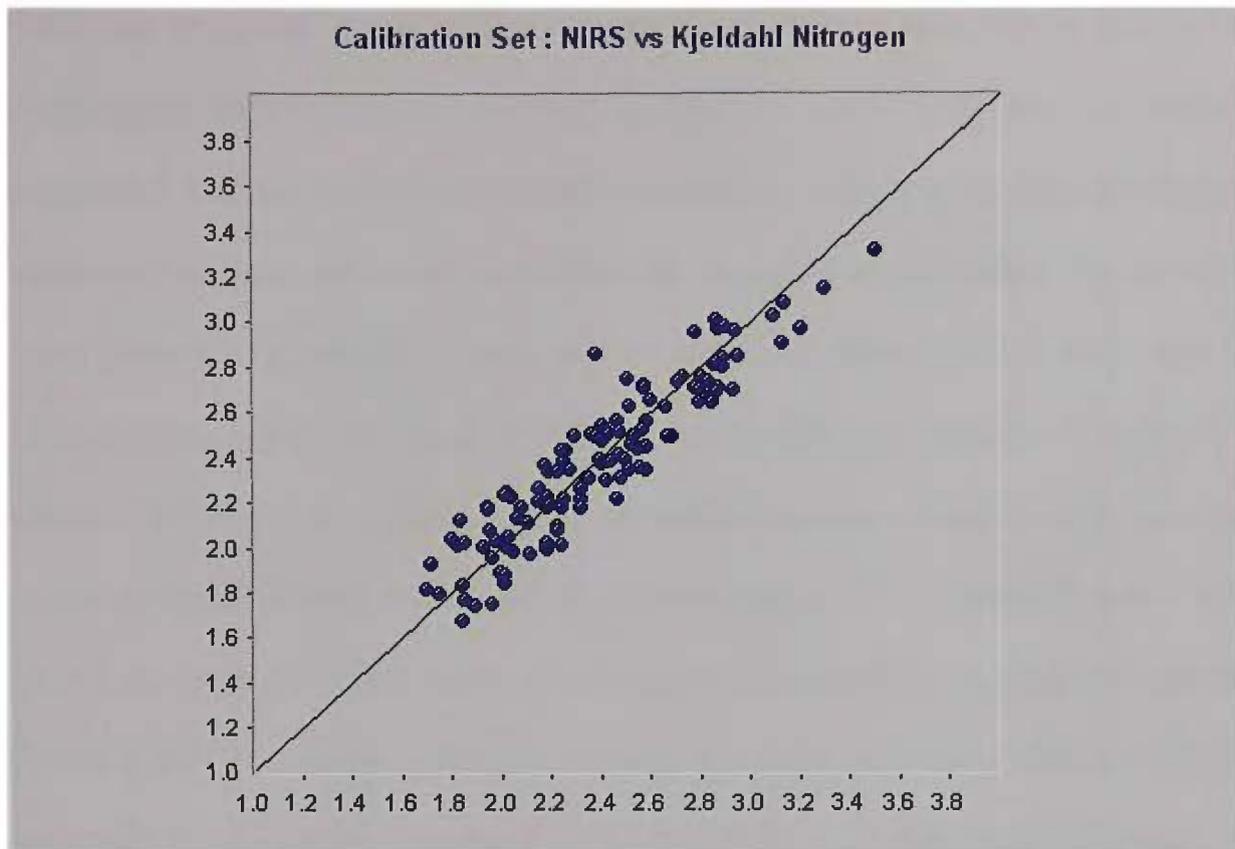


Figure 5.23 A plot of the combined calibration for Kjeldahl protein using PLS

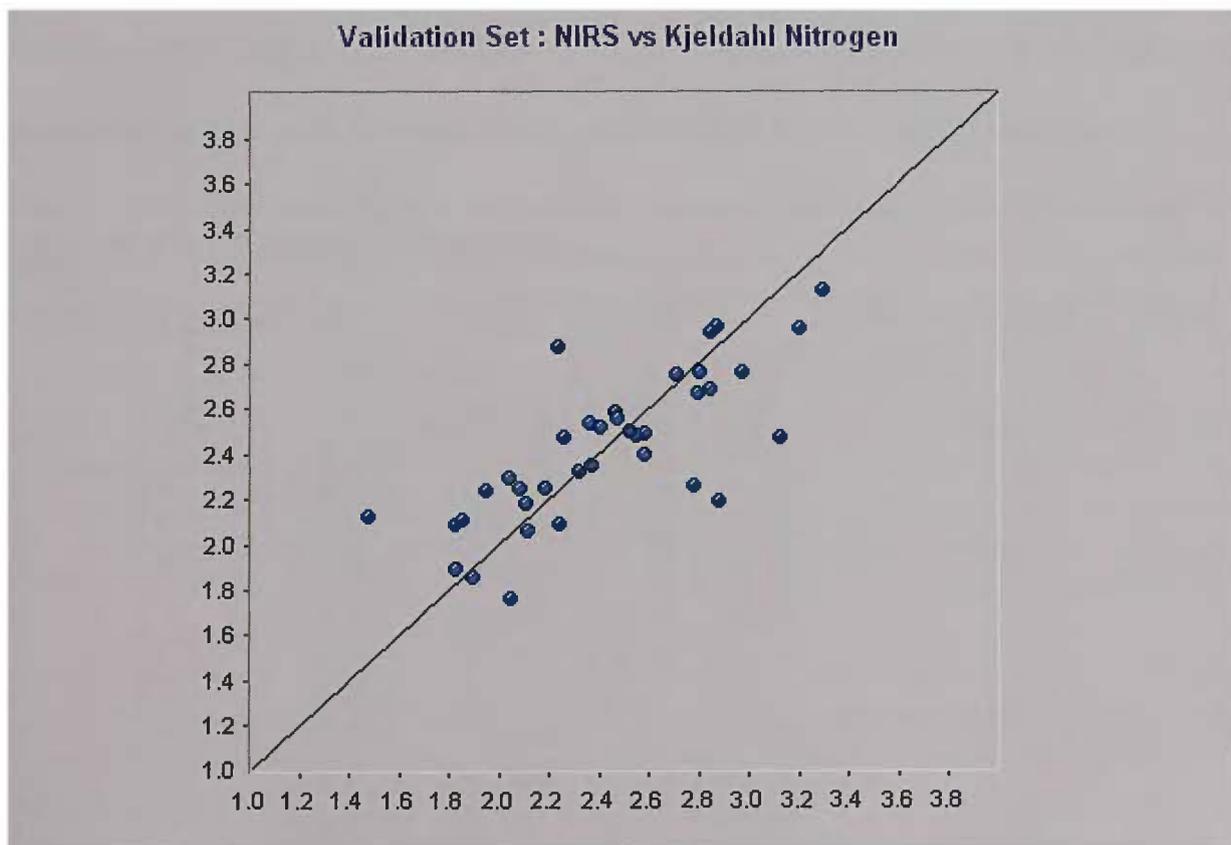


Figure 5.24 A plot of the combined validation set for protein using PLS

Plots of the PLS calibration set (Figure 5.23) and the validation set (Figure 5.24) are provided above. Kjeldahl protein of sultanas was one of the more challenging calibrations to develop, as can be seen from the correlation coefficient that was 0.793 for the MLR calibration, and the number of factors used and outliers removed to achieve an acceptable calibration. As can be seen from the correlation coefficient of both calibrations, PLS achieved a considerably higher  $R^2$  value of 0.859 (Table 5.24) as compared to the  $R^2$  value of 0.793 for MLR (Table 5.23), this difference was more notable than in other constituent calibrations, which indicate that MLR calibrations were not as robust as PLS, in this work. The F-value obtained by the PLS calibration (Table 5.24 F-value=44.4) was lower than the MLR calibration (Table 5.23 F-value=72.1), F-values achieved with the individual seasons' calibrations (Table 5.15 Season 1 PLS F-values=35.3; Table 5.18 Season 2 PLS F-values= 58.1; Table 5.21 Season 3 PLS F-values=35.5) were in one case of comparable size and showed improvement over the first and last seasons.

Table 5.24 Factors contributing to the combined seasons' calibration of Kjeldahl protein (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

PLS								
Factors	Outliers	$R^2$ (Cal)	SEC	SECV	F- value	$R^2$ (Val)	Bias	SEP
1	15	0.039	0.377	0.381	4.69	-	-	-
2	"	0.139	0.358	0.372	9.21	-	-	-
3	"	0.335	0.316	0.359	19.02	-	-	-
4	"	0.371	0.309	0.350	16.52	-	-	-
5	"	0.479	0.283	0.338	20.40	-	-	-
6	"	0.551	0.263	0.359	22.51	-	-	-
7	"	0.584	0.255	0.334	21.88	-	-	-
8	"	0.635	0.240	0.341	23.51	-	-	-
9	"	0.701	0.218	0.335	27.93	-	-	-
10	"	0.735	0.206	0.332	29.47	-	-	-
11	"	0.762	0.196	0.324	30.62	-	-	-
12	"	0.820	0.172	0.313	39.36	-	-	-
13	"	0.836	0.164	0.315	40.52	-	-	-
<b>14</b>	<b>15</b>	<b>0.859</b>	<b>0.153</b>	<b>0.307</b>	<b>44.43</b>	<b>0.758</b>	<b>0.0226</b>	<b>0.316</b>

The calibrations F-value also showed a clear improvement as factors were added to the calibration (Table 5.24 Factor=1 F-value=4.69; Factor=14 F-value=44.43), which indicated that the calibration was robust [8]. As the calibration developed by the addition of more factors both the coefficient of determination and SEC improved sharply (Table 5.24 Factor=1  $R^2=0.039$ , SEC=0.377; Factor=14  $R^2=0.859$ , SEC=0.147). Error of cross validation has exhibited a steady improvement as more factors were added to the PLS calibration (Table 5.24 Factor=1 SECV=0.381; Factor=14 SECV=0.307), meaning that the process of optimisation has not over-fitted the calibration and this supports the trend seen within the F-value [9]. The coefficient of determination of the validation was 0.758 with an uncorrected SEP of 0.316 and Bias of 0.0226 (Table 5.24).

### 5.3.5 Comparison to cited literature

Flinn *et al* [13] investigated the ability of diffuse reflectance NIRS to assess a series of chemical constituents and processing characteristics (including crude protein) of both ground and whole samples of pulses, more specifically chickpeas and field peas.

Comparing the calibration of the first seasons' samples for Kjeldahl protein (KP), it can be seen that its performance is not as robust to the calibrations of whole chickpeas and field peas by Flinn *et al*. Whole Chickpeas achieved a coefficient of determination of 0.94 and an error of cross validation of 0.57. The whole field pea calibration also proved to have a higher coefficient of determination than that developed from the first seasons' samples of sultanas (Flinn *et al* Field peas  $R^2=0.92$ , SECV=0.58; Table 5.15 PLS Season 1 sultanas  $R^2=0.85$ , SECV=0.33). This result was not surprising due to the

nature of the sultana sample set, as there were only a small number of samples from this season, too small to be considered a wide-ranging and representative series of samples [13]. The calibration from the second seasons samples, in particular PLS compares more closely with the calibrations by Flinn and co workers, achieving a correlation coefficient of 0.898 and error of cross validation of 0.194 (Table 5.17) which is close to the correlation coefficient of the field pea calibration developed by Flinn *et al* (Flinn *et al* field peas  $R^2=0.92$  SECV=0.58) [13].

The third seasons' calibration was not as robust as calibrations by Flinn and co workers of whole chickpeas and field peas. In particular the MLR calibration was not as reliable since the coefficient of determination of 0.765 (Table 5.20). PLS was also not as robust when developing a regression with this data set, yielding a coefficient of determination of 0.840 (Table 5.21). This is considerably less than the calibrations developed by Flinn *et al* for whole pulses (Flinn *et al* Whole field peas  $R^2=0.92$  SECV=0.58; Whole chickpeas  $R^2=0.94$  SECV=0.57). In addition, the sultana data set contained a number of redundant samples, samples that contribute very little to the variability of the calibration [13].

Compared to the calibrations developed by Flinn *et al*, the combined seasons' calibrations were not as robust, which is clearly indicated by the coefficients of determination (Combined KP calibrations: Table 5.23 MLR  $R^2=0.79$ ; Table 5.24 PLS  $R^2=0.86$ ), which are significantly lower than the coefficient of determination of the pulse calibrations for crude proteins (Flinn *et al* whole field peas  $R^2=0.92$ ; whole chickpeas  $R^2=0.94$ ) [13]. The low coefficient of determination for the sultana calibrations may in part be due to the

comparative uniformity of pulses compared to sultanas, also the relative transparency and compact-ability of sultanas compared to whole pulses may also effect the comparative coefficient of determination of these calibrations, despite the similarity in spectroscopic technique and physical size of both sultanas and pulses however, the very different texture and density of the two products is a very important point to take into consideration. Two considerably different agricultural products are being compared in this discussion.

Similarly, the calibration developed by Fassio and Cozzolino [16] for the composition determination of whole sunflower seeds obtained a coefficient of determination of 0.96, which is clearly more robust calibration than that developed for sultanas (Combined seasons KP calibration: Table 5.23 MLR  $R^2=0.79$ ; Table 5.24 PLS  $R^2=0.86$ ). This is also likely to be due to the uniform nature of sunflower seeds compared to sultanas, their texture and constant packing characteristics.

Confalonieri *et al* [17] used reflectance NIRS spectroscopy to asses natural alpine swards for dry matter constituents, in particular crude protein yielding a coefficient of determination of 0.97 using modified PLS utilising software elimination. Kays and co-workers [18] utilised reflectance NIRS to asses constituents of diverse cereal grain food products simultaneously with one calibration achieving a coefficient of determination of 0.973 for protein A (nitrogen x 6.25). White *et al* [19] used NIRS to asses wet diet of commercial minks chemical constituents, including crude protein which achieved a coefficient of determination of 0.96. Comparing the combined calibrations of sultanas (Combined seasons KP calibration: Table 5.23 MLR  $R^2=0.79$ ; Table

5.24 PLS  $R^2=0.86$ ) to the three projects above it is clear that the nature of the analyte being studied affects the coefficient of determination obtained, as each case was a highly homogeneous medium and therefore achieved a more strongly correlating calibration.

## 5.4 Titratable Acidity

### 5.4.1 Season one titratable acidity

Table 5.25 Laboratory parameters summary of season one TA

# Samples (in calibration)	# Samples in validation	Range	Mean	Standard Deviation
35(28)	7	199-281	225	17.11

#### 5.4.1.1 MLR Calibration

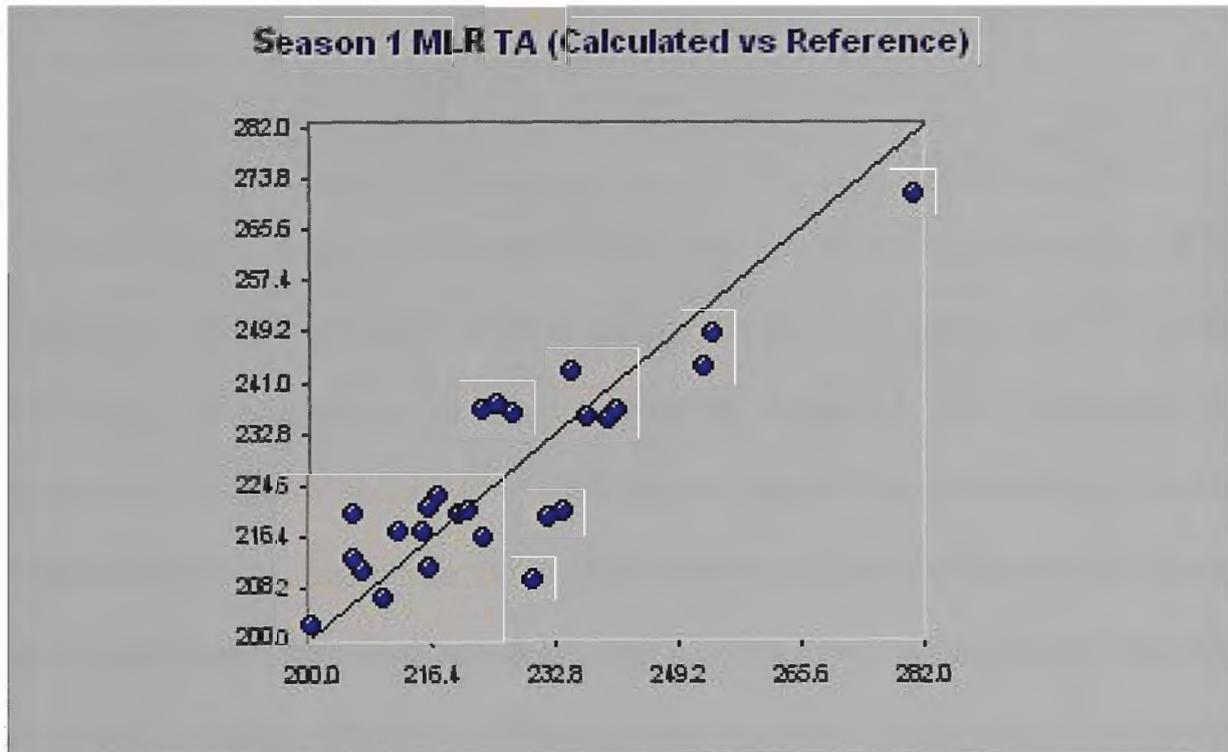


Figure 5.25 A plot of the first seasons calibration for titratable acidity using MLR

The first term (Table 5.26 1056nm) used to develop the calibration for titratable acidity was due to stretching and deformation vibrations of carboxylic, aryl and alkyl hydroxy groups, as well as amines and amides, also carbon hydrogen bonds of alkyl groups and aldehydes absorb at this wavelength. The second term (1224nm) was attributed to the stretching vibrations of carbon hydrogen bonds [5] of alkyl groups and alkenyl groups with a single hydrogen bonded [6]. The third term (616nm) comes from the visible region [7], which is reasonable as titratable acidity can be linked to the maturity and “green-ness” of fruit [20]. The fourth term (2042nm) is linked to

symmetrical and asymmetrical stretching vibrations [5], typically seen when proteins and amines, aqueous ammonia and secondary and tertiary amides are present [7], also terminal alkenyl groups and alkyl and carboxylic acid hydroxy groups as well as water absorb at this wavelength [6].

Table 5.26 wavelengths contributing to the first seasons MLR calibration for titratable acidity (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

MLR					
Term	Outliers	$\lambda$ (nm)	R <sup>2</sup> (Cal)	SEC	F-value
1	-	1056	0.281	15.53	9.37
2	-	1224	0.504	13.18	11.7
3	-	616	0.673	10.94	15.1
4	-	<b>2042</b>	<b>0.763</b>	<b>9.54</b>	<b>16.8</b>

Given above is a plot of the calibration set for MLR (Figure 5.25). The coefficient of determination of this calibration is low (Table 5.26 R<sup>2</sup>=0.76, SEC=9.5) compared to other constituents, however the calibration is preliminary in nature and small in size, 28 samples in the calibration set, such limitations will be overcome once the calibrations are combined. As more terms were added to this calibration the coefficient of determination steadily improved, a steady increase in F-value was also seen indicative of a reliable calibration (Table 5.26 Term=1 R<sup>2</sup>=0.281, SEC=15.53, F-value=9.37; Term=4 R<sup>2</sup>=0.763, SEC=9.54, F-value=16.8) [8]. The size of the F-value gives a clear indication that the calibration is only a preliminary one that is based on 35 sultana samples for one season but is giving a good indication that a larger sample number would achieve a considerably higher correlation [8].

### 5.4.1.2 PLS Calibration

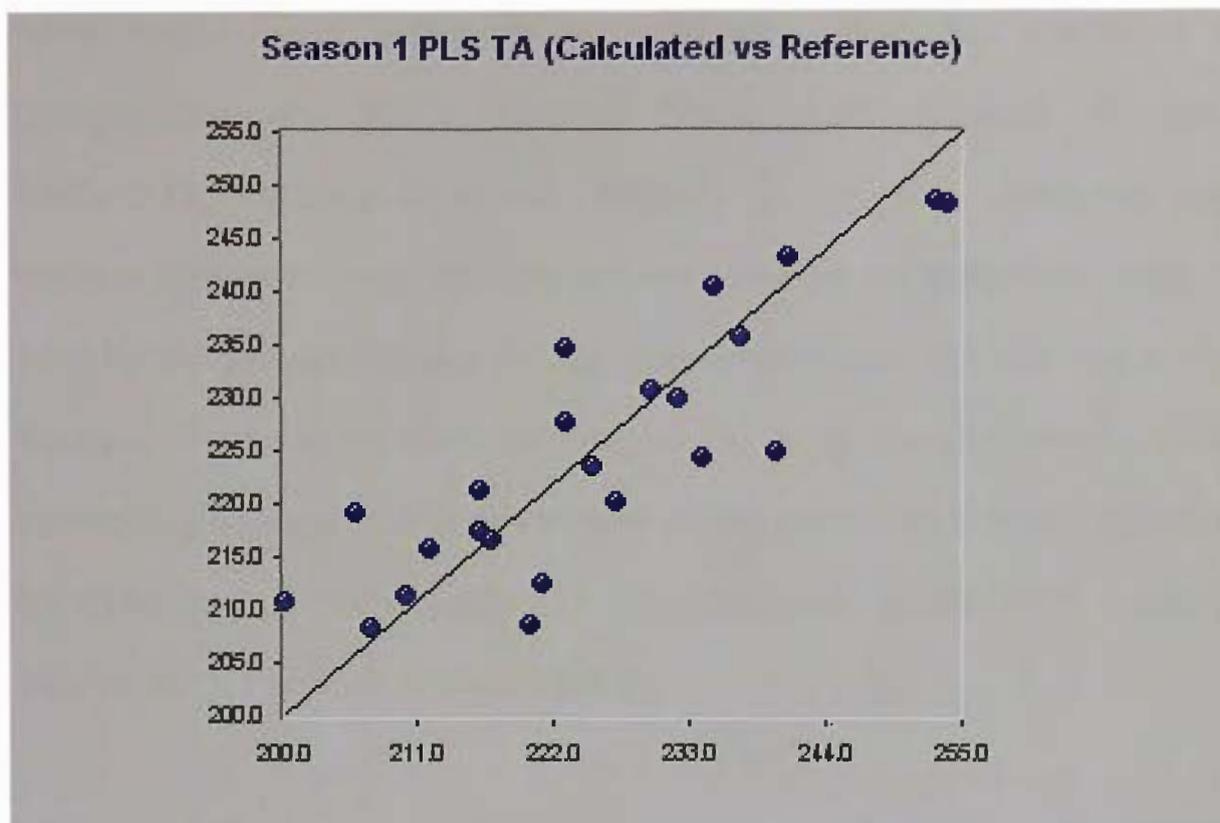


Figure 5.26 A plot of the first seasons' calibration for titratable acidity using PLS

Above is a plot of the PLS calibration (Figure 5.26). The calibrations developed from the titratable acidity results from the first season are different to calibrations of other constituents, MLR has shown an increased correlation coefficient and decreased error of calibration compared to PLS (Table 5.26 MLR  $R^2=0.763$ ,  $SEC=9.54$ ,  $F\text{-value}=16.8$ ; Table 5.27 PLS  $R^2=0.741$ ,  $SEC=7.81$ ,  $F\text{-value}=18.17$ ), though the F-value of the PLS calibration is slightly higher than the MLR calibration.

Table 5.27 Factors contributing to the first seasons' PLS calibration of titratable acidity (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

PLS					
Factors	Outliers	$R^2$ (Cal)	SEC	SECV	F-value
1	3	0.244	12.71	13.71	6.78
2	3	0.591	9.58	12.33	14.46
<b>3</b>	<b>3</b>	<b>0.741</b>	<b>7.81</b>	<b>11.03</b>	<b>18.17</b>

This atypical result can be attributed to the small size of the calibrations. As more factors were added to the calibration, both the coefficient of determination and SEC improved (Table 5.27 Factor=1  $R^2=0.244$ , SEC=12.71; Factor=3  $R^2=0.741$ , SEC=7.81). The PLS calibration also displays a lower F-value, that only indicates that the calibration was valid, as seen by the general increase of F as more factors were included (Table 5.27 Factor=1 F-value=6.78; Factor=3 F-value=18.17) [8]. Standard error of cross validation also supports this observation as this parameter steadily decreases as more factors were added to this calibration (Table 5.27 Factor=1 SECV=13.71; Factor=3 SECV=11.03) [9].

## 5.4.2 Season two titratable acidity

Table 5.28 Laboratory parameters summary of season two TA

# Samples (in calibration)	# Samples in validation	Range	Mean	Standard Deviation
95 (76)	19	148.1-336.7	217.8	24.39

### 5.4.2.1 MLR calibration

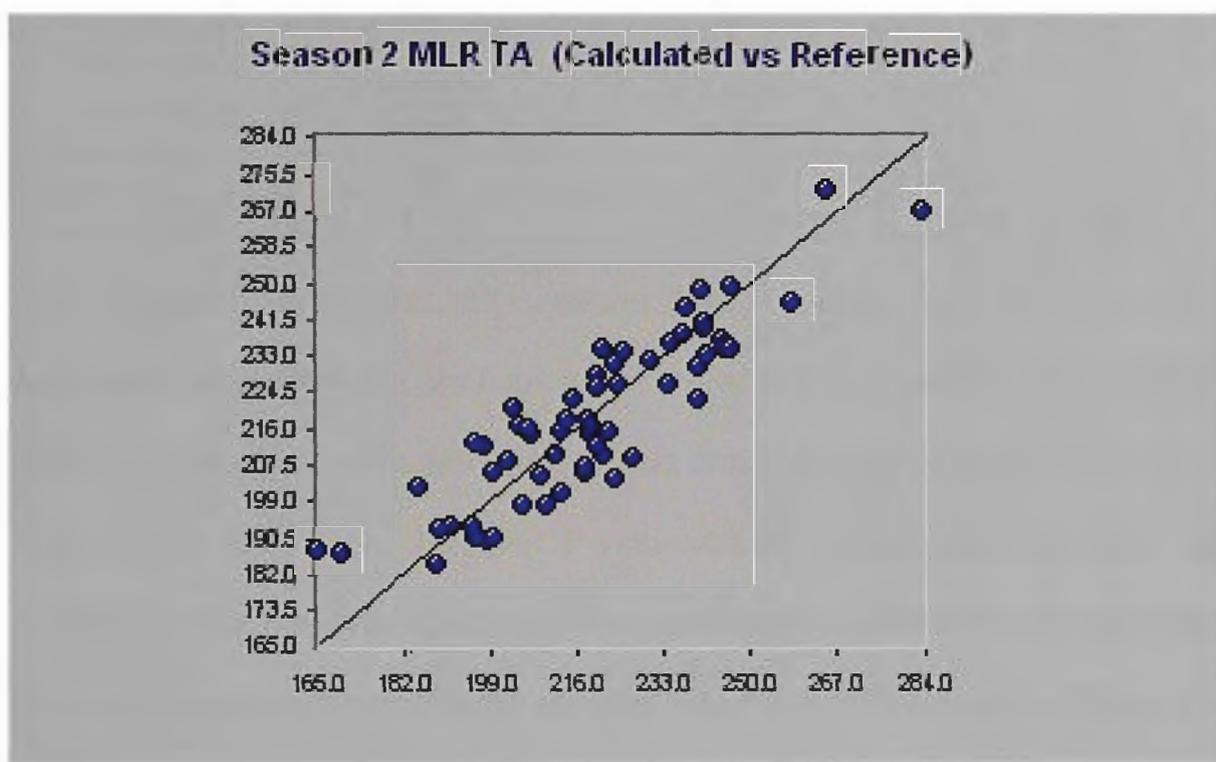


Figure 5.27 A plot of the second seasons' calibration for titratable acidity using MLR

The first term (Table 5.29 2222nm) of this calibration can be attributed to the stretching and deformation vibrations of nitrogen hydrogen bonds [5], typically seen when organic ammonium ions are present, often linked to amino acids, also methyl groups, secondary amides, secondary amines also absorb at this wavelength [6]. The second term (782nm) came from a region of the spectra associated with the stretching vibrations of nitrogen-hydrogen bonds [5], typically seen when primary amines are present [6]. The third term (1260nm) originated from a region close to the 1<sup>st</sup> overtone combinations band [5]. The fourth term (1132nm) arose from the second overtone stretching vibrations of

carbon-hydrogen bonds [5] of aryl groups, primary and tertiary alkanes, alkenes (except terminal alkenes) and aldehydes, as well as primary and secondary amides [6]. A plot of the MLR calibration is given above (Figure 5.27).

Table 5.29 Wavelengths contributing to the second seasons' MLR calibration of titratable acidity (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

MLR					
Term	Outliers	$\lambda$ (nm)	R <sup>2</sup> (Cal)	SEC	F-value
1	8	2222	0.4092	17.1	40.9
2	"	782	0.6344	13.6	50.3
3	"	1260	0.748	11.4	56.3
<b>4</b>	<b>8</b>	<b>1132</b>	<b>0.794</b>	<b>10.4</b>	<b>53.9</b>

The calibration for the second seasons F-value (Table 5.29 F-value= 53.9) was much larger than the previous season (Table 5.26 F-value=16.8), and the trend seen as more terms were added was one of steady increase (Table 5.29 Term=1 F-value=40.9; Term=4 F-value=53.9), which showed that the calibration was not over-fitted [8]. The coefficient of determination and SEC improved steadily as more terms were included into the calibration (Table 5.29 Term=1 R<sup>2</sup>=0.409, SEC=17.4; Term=4 R<sup>2</sup>=0.79, SEC=10.4). The calibration of the second season showed a marked improvement over the MLR calibration of the first season, due largely to the increase in size of the second seasons calibration over the first, resulting from a more representative calibration. This improvement in the calibration has resulted in an increased coefficient of determination (Table 5.26 Season 1 R<sup>2</sup>=0.76, SEC=9.5; Table 5.29 Season 2 R<sup>2</sup>=0.79, SEC=10.4) over the first season.

### 5.4.2.2 PLS calibration

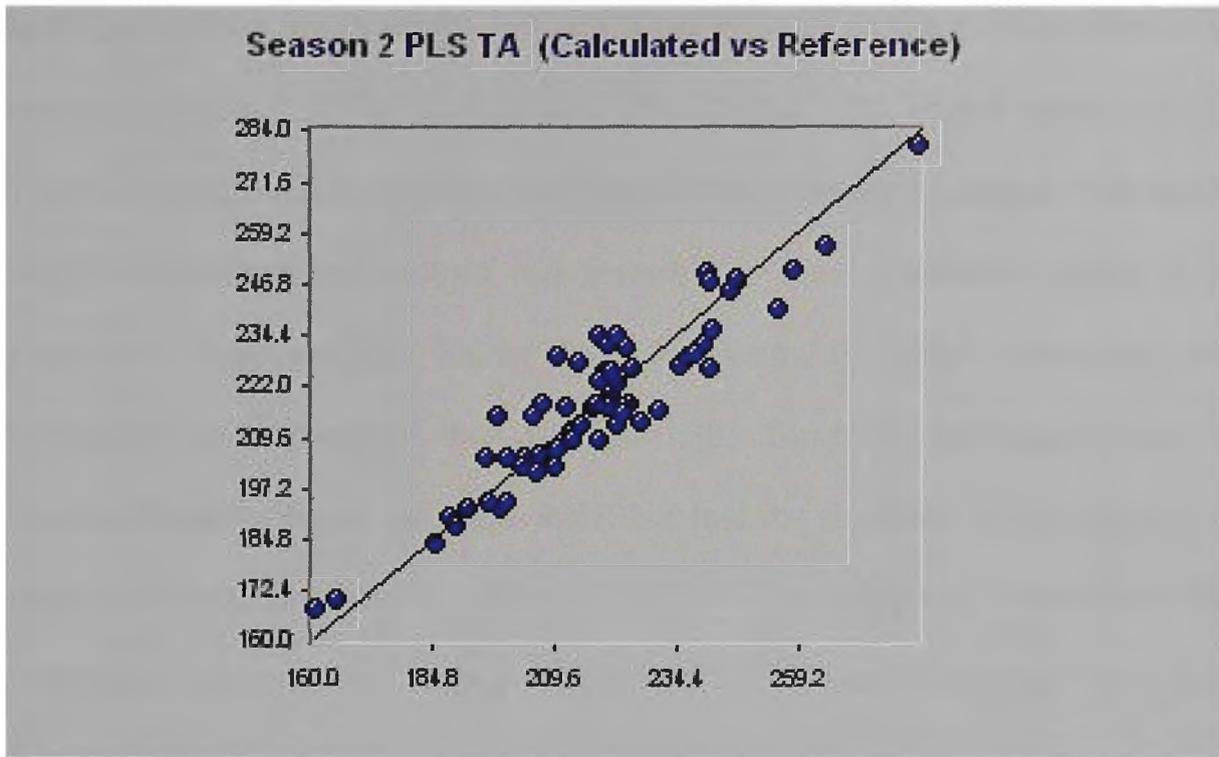


Figure 5.28 A plot of the second seasons' calibration for titratable acidity using PLS

Given above is a plot of the PLS calibration (Figure 5.28) of titratable acidity. The second seasons' calibrations showed performance more typical than the previous season, PLS gave a distinctly higher coefficient of determination than MLR (Table 5.29 MLR  $R^2=0.79$ ,  $SEC=10.4$ ; Table 5.30 PLS  $R^2=0.87$ ,  $SEC=8.49$ ).

Table 5.30 Factors contributing to the second seasons' PLS calibration of titratable acidity (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

PLS					
Factors	Outliers	$R^2$ (Cal)	SEC	SECV	F-value
1	8	0.082	21.18	21.48	5.37
2	"	0.444	16.61	17.69	23.60
3	"	0.593	14.34	16.78	28.20
4	"	0.654	13.34	16.31	26.92
5	"	0.778	10.79	14.91	39.15
6	"	0.809	10.08	14.29	38.91
7	"	0.836	9.42	14.30	39.47
<b>8</b>	<b>8</b>	<b>0.870</b>	<b>8.49</b>	<b>15.12</b>	<b>44.23</b>

As more terms were added to the calibration both the error of calibration and coefficient of determination improved sharply (Table 5.30 Factor=1  $R^2=0.082$ ,

SEC=21.18; Factor=8  $R^2=0.87$ , SEC=8.49). The difference in the F-values of both calibrations was notable but not large enough to affect the calibrations reliability (Table 5.29 MLR: F-value= 53.9; Table 5.30 PLS F-value= 44.2). The PLS calibration for the second season also showed a general increase in the F-value as more factors are included in the calibration (Table 5.30 Factor=1 F-value=5.37; Factor=8 F-value=44.23) which indicated the increasing robustness of the calibration [8]. Similarly, the trend seen in standard error of cross validation indicated that the reliability of this calibration had not been significantly affected by the introduction of factors to this calibration, as an overall steady decrease in SECV was observed (Table 5.30 Factor=1 SECV=21.48; Factor=8 SECV=15.12) [9].

### 5.4.3 Season three Titratable Acidity

Table 5.31 Laboratory parameters summary of season three TA

# Samples (in calibration)	# Samples in validation	Range	Mean	Standard Deviation
156(125)	31	167.5-359.3	211.6	26.5

#### 5.4.3.1 MLR calibration

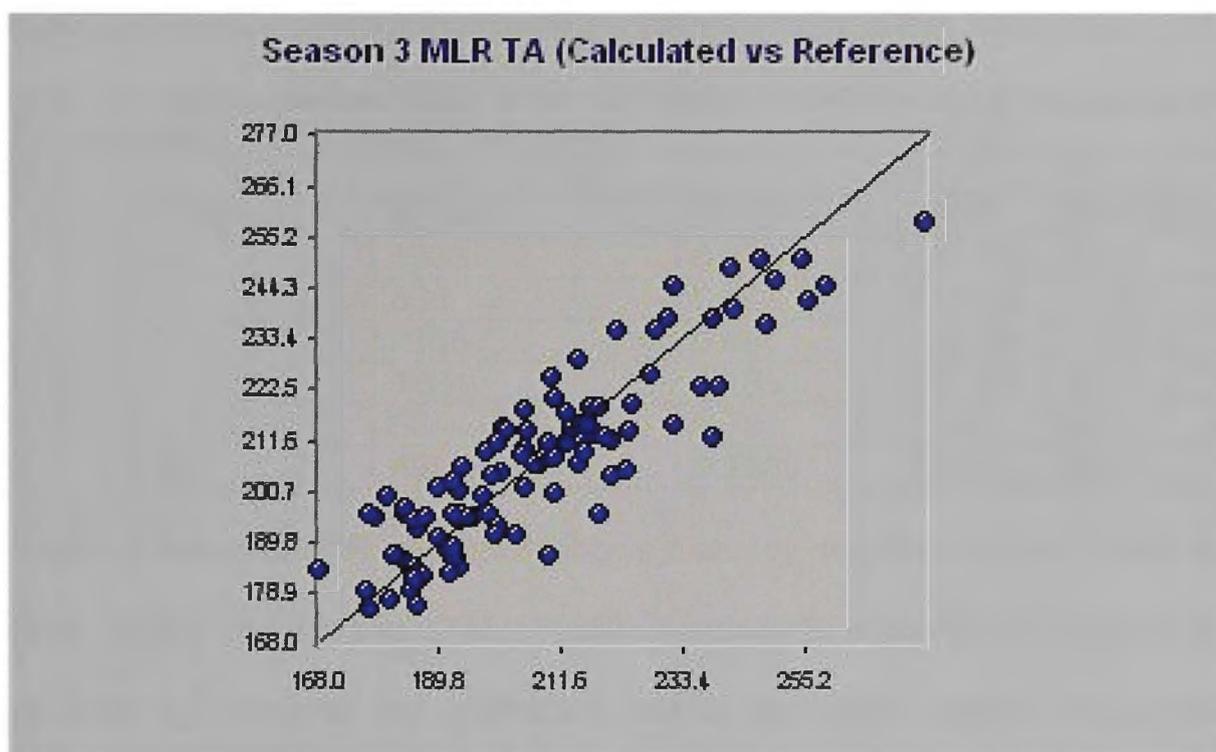


Figure 5.29 A plot of the third seasons' calibration for titratable acidity using MLR

The first summation term (Table 5.32 1966nm) of the calibration of titratable acidity was due to tertiary, aryl and carboxylic acid hydroxy groups and water absorb at this wavelength [6], as well as the asymmetric stretching vibrations of nitrogen hydrogen bonds [5] in primary and secondary amides, primary amines and aqueous ammonia, also ketones [6]. The second term (658nm) originated from the visible region of the spectrum [7]. The third term (2136nm) was due to the stretching vibrations of nitrogen hydrogen bonds and carbonyl double bonds [5] of amino acids [7] as well as cis and trans alkenyl and aryl carbon hydrogen bonds, alkyl and carboxylic acid hydroxy groups,

deprotonated carboxylates, esters, ketones, anhydrides all absorb at this wavelength [6]. Second overtone carbon hydrogen bond stretching vibrations of alkanes and alkenyl groups containing a single carbon hydrogen bond are the source of the fourth term (1224nm) [5]. The fifth and sixth terms (524nm and 492nm) have origins within the visible region of the spectrum [6].

A plot of the MLR calibration is provided above (Figure 5.29).

Table 5.32 Wavelengths contributing to the third seasons' MLR calibration of titratable acidity (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

MLR					
Term	Outliers	$\lambda$ (nm)	R <sup>2</sup> (Cal)	SEC	F-value
1	13	1966	0.466	15.73	87.4
2	"	658	0.591	13.84	71.5
3	"	2136	0.691	12.09	73.1
4	"	1224	0.735	11.25	67.3
5	"	524	0.781	10.3	68.3
<b>6</b>	<b>13</b>	<b>492</b>	<b>0.7965</b>	<b>9.97</b>	<b>62.0</b>

The third seasons MLR calibration shows a steady decrease in the F-value as more terms are added (Table 5.32 Term=1 F-value=87.4; Term=6 F-value=62.0), however the calibration has a sufficiently large F-value that would indicate a capacity to make reliable predictions. With the introduction of more terms to the calibration both the coefficient of determination and SEC improved (Table 5.32 Term=1 R<sup>2</sup>=0.466, SEC=15.73; Term=6 R<sup>2</sup>=0.797, SEC=10.0). The coefficient of determination achieved by this calibration is acceptable considering this calibrations preliminary nature, and is higher than previous MLR calibrations for this constituent (Table 5.26 Season 1 R<sup>2</sup>=0.763, SEC=9.54; Table 5.29 Season 2 R<sup>2</sup>=0.79, SEC=10.4; Table 5.32 Season 3 R<sup>2</sup>=0.797, SEC=10.0).

### 5.4.3.2 PLS Calibration

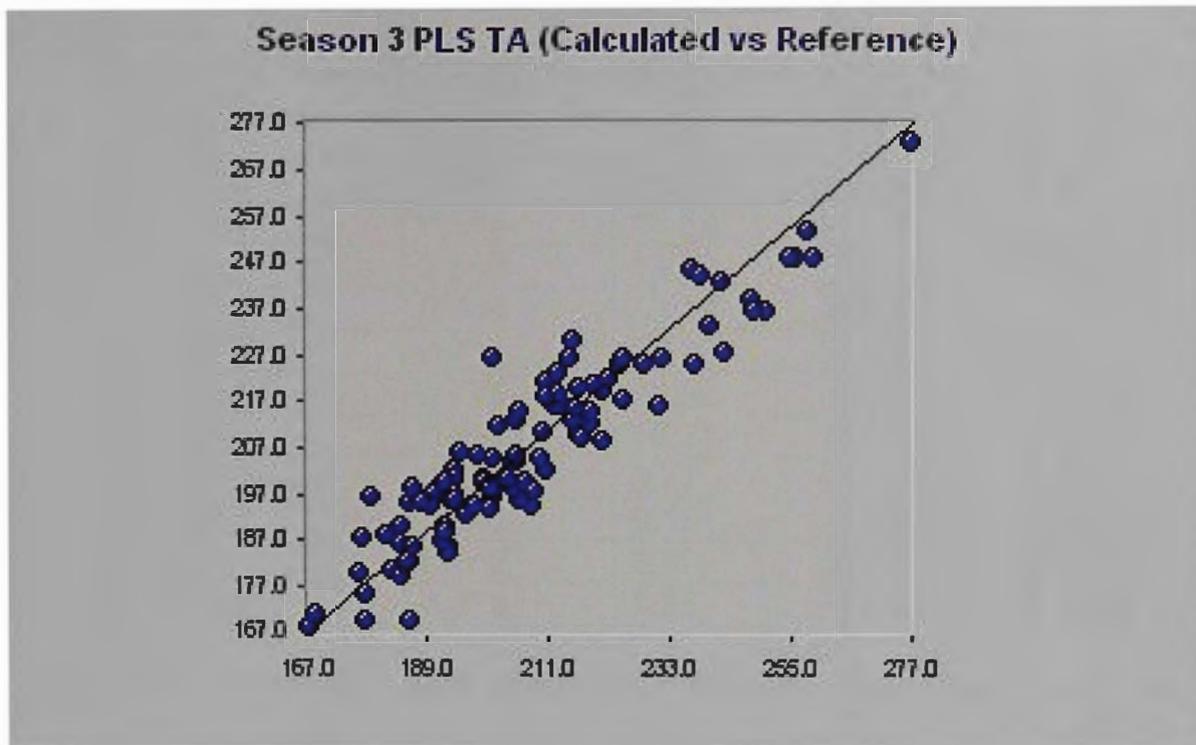


Figure 5.30 A plot of the first seasons' calibration for titratable acidity using PLS

Provided above is a plot of the PLS calibration set (Figure 5.30). The calibrations of the third seasons' titratable acidity results show behaviour typically seen throughout this study, PLS is in almost all cases developed higher correlating calibrations than MLR (Table 5.32 MLR  $R^2=0.797$ ; Table 5.33 PLS  $R^2=0.878$ ). The coefficient of determination was higher than that achieved by the MLR calibration, and the F-value of the PLS Calibration (Table 5.32 MLR F-value=62.0; Table 5.33 PLS F-value= 48.6) was less, but this was not notable. As more factors were included into the calibration both the coefficient of determination and SEC improved (Table 5.33 Factor=1  $R^2=0.177$ , SEC=20.1; Factor=13  $R^2=0.878$ , SEC=8.26). The PLS calibration showed a trend in the F-value of a steady increase as more factors were include in the calibration (Table 5.33 Factor=1 F-value=21.53; Factor=13 F-value=48.58), which showed that the calibration was not over-fitted [8]. The F-value itself is sufficiently large enough to consider this calibration robust,

Table 5.33 Factors contributing to the third seasons' PLS calibration of titratable acidity (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

PLS					
Factors	Outliers	R <sup>2</sup> (Cal)	SEC	SECV	F-value
1	13	0.177	20.10	20.42	21.53
2	"	0.508	15.62	16.76	51.09
3	"	0.545	15.10	16.23	39.12
4	"	0.584	14.52	15.87	34.03
5	"	0.643	13.52	16.15	34.53
6	"	0.664	13.18	16.19	31.27
7	"	0.711	12.29	16.30	33.02
8	"	0.775	10.89	16.11	40.13
9	"	0.801	10.31	15.84	41.07
10	"	0.823	9.77	15.45	42.32
11	"	0.847	9.13	14.97	45.41
12	"	0.862	8.72	14.83	46.40
<b>13</b>	<b>13</b>	<b>0.878</b>	<b>8.26</b>	<b>14.85</b>	<b>48.58</b>

despite the preliminary nature of this calibration [8]. The trends seen in error of cross-validation displayed a steady improvement as more factors were added to this calibration (Table 5.33 Factor=1 SECV=20.42; Factor=13 SECV=14.85), which indicates an increase in the robustness of the calibrations predictions as it was optimised [9].

### 5.4.4 Combined seasons titratable acidity

Table 5.34 Laboratory parameters summary of combined TA

# Samples (in calibration)	# Samples in validation	Range	Mean	Standard Deviation
184 (147)	37	148.1-359.3	216.2	29.79

#### 5.4.4.1 MLR Calibration

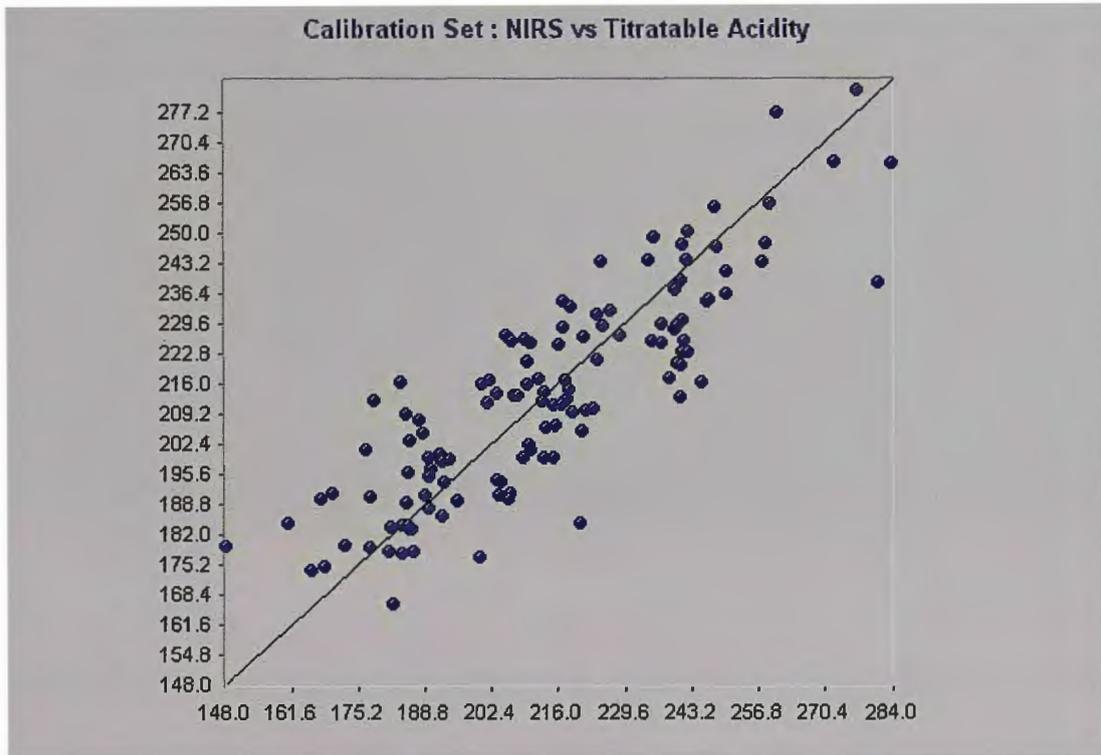


Figure 5.31 A plot of the combined calibration for titratable acidity using MLR

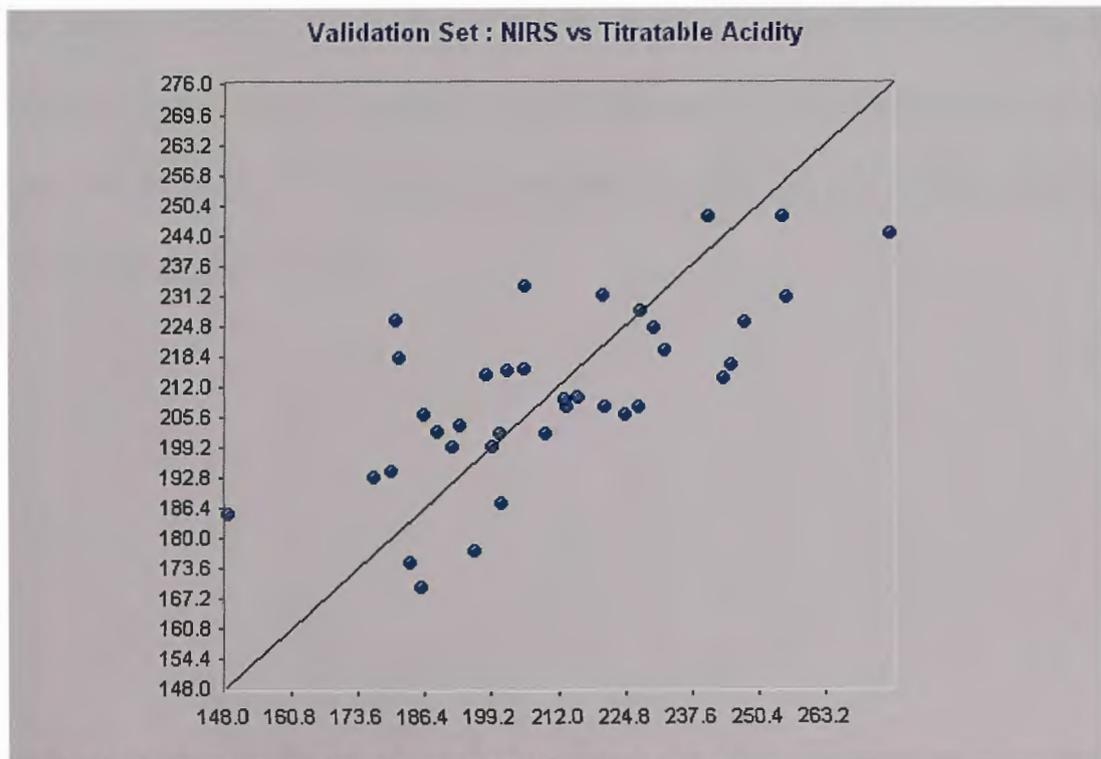


Figure 5.32 The combined validation set for titratable acidity using MLR

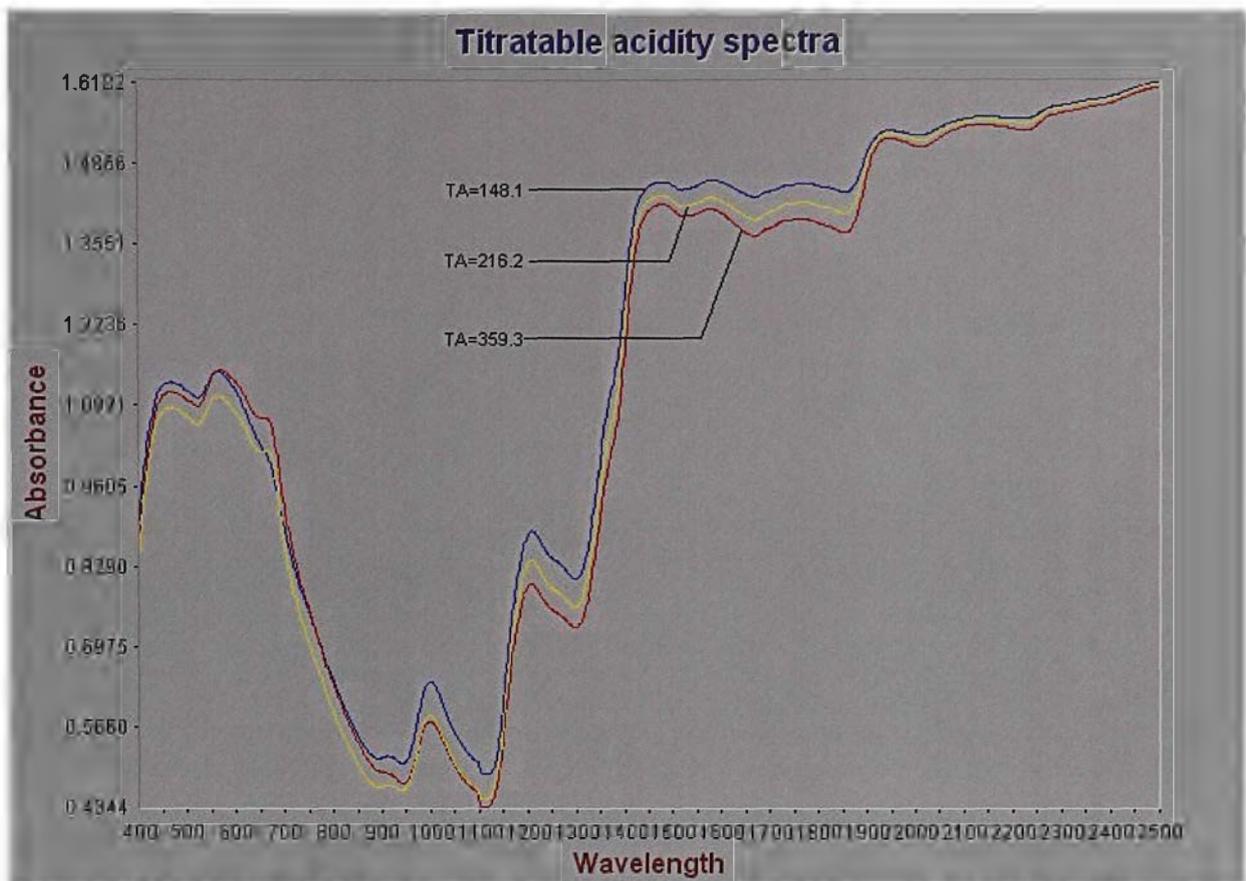


Figure 5.33 Spectra of the minimum (blue), maximum (red) and mean (yellow) samples of titratable acidity

A diagram containing three spectra of the titratable acidity calibration is shown above (Figure 5.33), the spectra of the sample with the highest titratable acidity value (359.3) in red, the spectra of the sample closest to the mean (216.4) in yellow and the spectra of the sample with the lowest value (148.1) in blue. A close up of the region of the first wavelength of the MLR calibration (Figure 5.34) is given below.

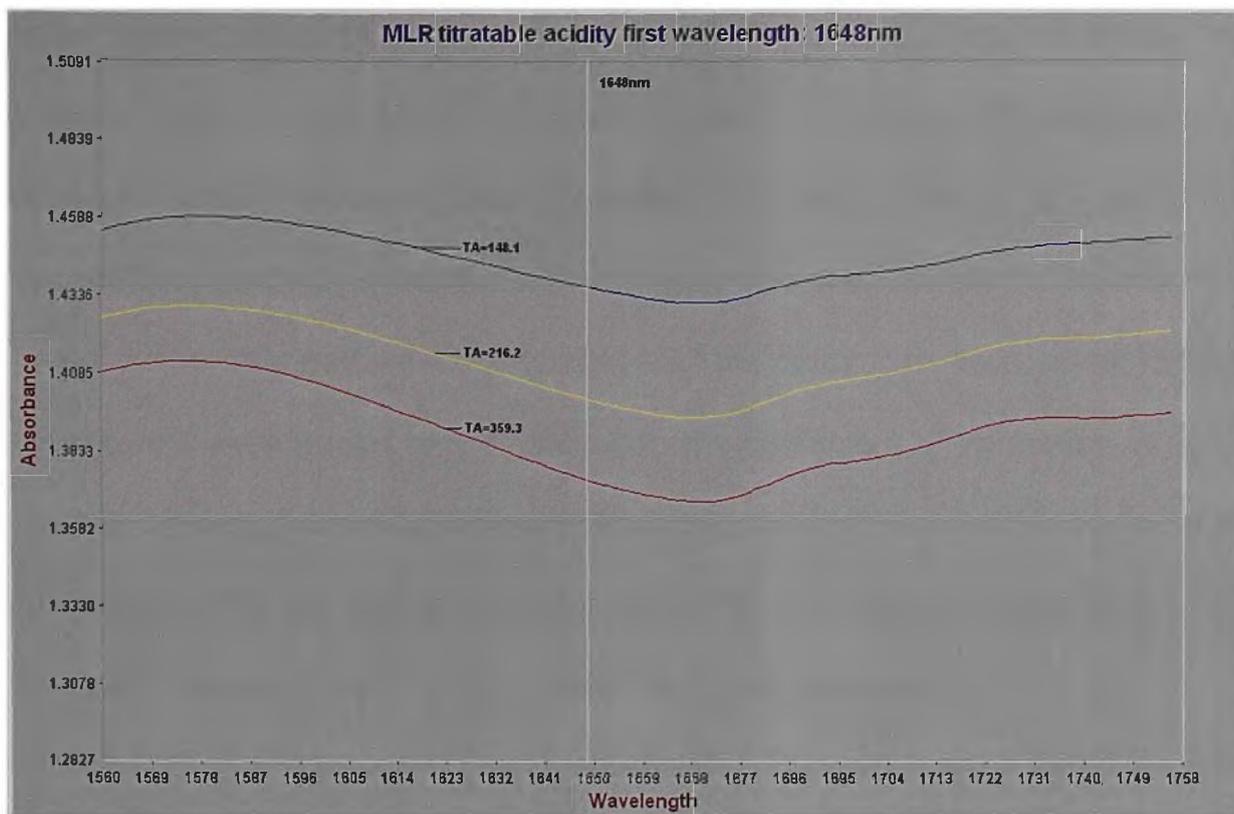


Figure 5.34 Close up of NIR spectra at the first wavelength of the combined MLR calibration. The first term (Table 5.35 1648nm) of the combined calibration of titratable acidity was linked to the carbon hydrogen bond stretching vibrations [5] linked to alkenes or aryl groups, also secondary amines absorb at this wavelength [6]. The second term (658nm) originated within the visible region of the spectra [7]. The third term (1258nm) originated from the first overtone of carbon hydrogen combinations band [5]. The fourth term (512nm) was from the visible region of the spectrum [7]. The fifth term (1550nm) was attributable to the stretching vibrations of oxygen hydrogen bonds [5] of hydroxy or water molecules, also amines and aqueous ammonia absorb at this wavelength [6]. The final term (1980nm) of this calibration can be traced to aryl, carboxylic acid, secondary and tertiary hydroxy groups [6] and the asymmetric stretching of nitrogen hydrogen bonds within proteins [7], also the C-H bond of aldehydes absorb in that region [6].

Given above are plots of the calibration set (Figure 5.31) and validation set (Figure 5.32) of the MLR calibration. The MLR combined calibration of titratable acidity showed a steady decrease in F-value (Table 5.35 Term=1 F-value=75.6; Term=6 F-value=53.3), however this was compensated with the final F-value obtained, which indicated that the calibration was robust [8]. As more terms were added to the calibration the coefficient of determination and SEC improved (Table 5.35 Term=1  $R^2=0.387$ , SEC=21.56; Term=6  $R^2=0.73$ , SEC=14.5), though the final coefficient of determination achieved by this calibration was of limited utility, capable of coarse screening.

Table 5.35 Wavelengths contributing to the combined MLR calibration of titratable acidity (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

MLR								
Term	Outliers	$\lambda$ (nm)	$R^2$ (Cal)	SEC	F-value	$R^2$ (Val)	Bias	SEP
1	15	1648	0.387	21.56	75.6	-	-	-
2	"	658	0.527	19.02	66.3	-	-	-
3	"	1258	0.572	18.16	52.6	-	-	-
4	"	512	0.620	17.18	47.8	-	-	-
5	"	1550	0.678	15.9	48.7	-	-	-
<b>6</b>	<b>15</b>	<b>1980</b>	<b>0.736</b>	<b>14.5</b>	<b>53.3</b>	<b>0.70</b>	<b>1.68</b>	<b>19.77</b>

The calibration successfully predicted the validation set to a comparable coefficient of determination (Table 5.35  $R^2$  validation= 0.70), which reinforces the conclusion that this calibration was capable of coarse screening. The validation achieved an uncorrected SEP of 19.77 and Bias of 1.68 (Table 5.36).

### 5.4.4.2 PLS calibration

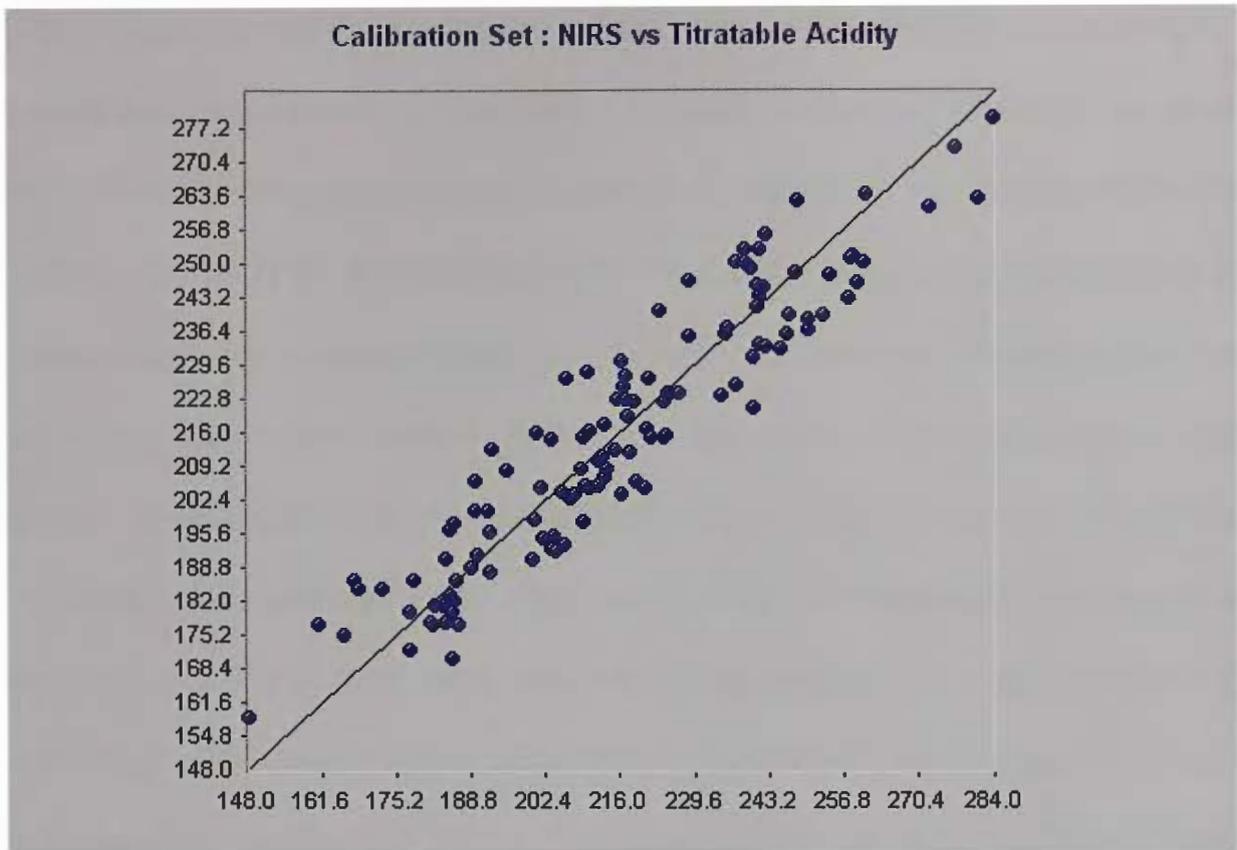


Figure 5.35 A plot of the combined calibration for titratable acidity using PLS

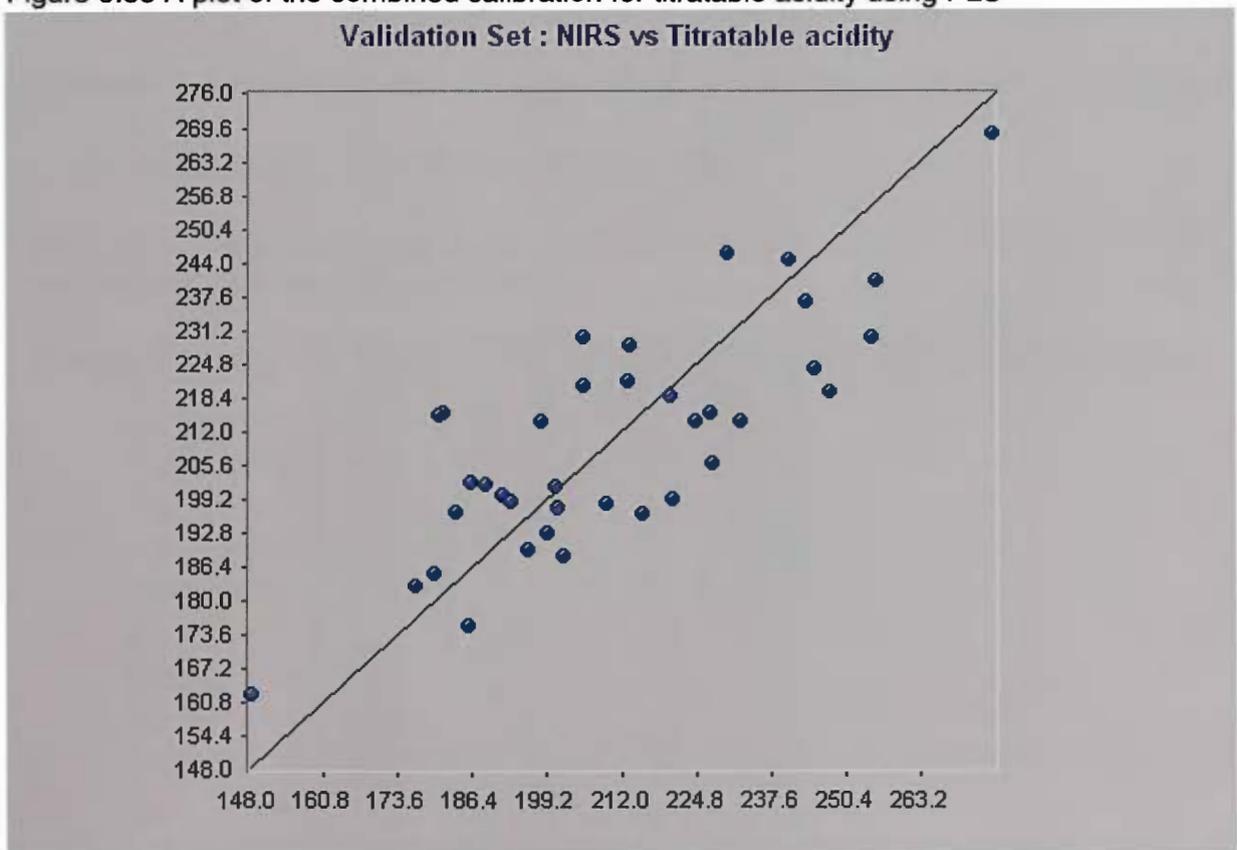


Figure 5.36 The combined validation set for titratable acidity using PLS

Above are provided plots of the calibration set (Figure 5.35) and validation set (Figure 5.36) of the PLS calibration. Titratable acidity was also a challenging constituent with respect to developing a NIRS calibration; MLR did not deal well with this constituent when compared to the other combined calibration (Table 5.35 MLR  $R^2=0.73$  SEC=14.5). This calibration would be considered only suitable for coarse screening, this was not the case, however, for the calibration developed using PLS that showed steady improvement as more factors were added to the calibration (Table 5.36 Factor=1  $R^2=0.179$ , SEC=25.1; Factor=15  $R^2=0.88$ , SEC=10.3). This PLS calibration required the removal of 10% outliers and the use of 15 factors, but has achieved a calibration with excellent performance for a calibration developed for a non-homogenous product by NIRS. The F-value of the PLS calibration was comparable to the corresponding MLR calibration (Table 5.35 MLR F-value=53.3; Table 5.36 PLS F-value=50.4), which showed the PLS calibration would prove reliable at predicting samples [8].

Table 5.36 Factors contributing to the combined PLS calibration of titratable acidity (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

PLS								
Factors	Outliers	$R^2$ (Cal)	SEC	SECV	F-value	$R^2$ (Val)	Bias	SEP
1	15	0.179	25.10	25.37	26.10	-	-	-
2	"	0.390	21.72	22.68	38.04	-	-	-
3	"	0.470	20.32	21.75	34.95	-	-	-
4	"	0.501	19.81	21.52	29.34	-	-	-
5	"	0.583	18.19	21.66	32.42	-	-	-
6	"	0.632	17.16	21.42	32.92	-	-	-
7	"	0.694	15.71	21.28	36.96	-	-	-
8	"	0.729	14.85	21.60	38.05	-	-	-
9	"	0.775	13.60	20.80	42.85	-	-	-
10	"	0.794	13.08	21.44	42.71	-	-	-
11	"	0.812	12.54	21.12	43.18	-	-	-
12	"	0.833	11.88	21.11	45.22	-	-	-
13	"	0.852	11.22	20.90	47.87	-	-	-
14	"	0.868	10.65	21.24	50.32	-	-	-
<b>15</b>	<b>15</b>	<b>0.877</b>	<b>10.33</b>	<b>21.34</b>	<b>50.40</b>	<b>0.798</b>	<b>-1.14</b>	<b>16.92</b>

The trend seen in the F-value with the addition of more factors to the PLS calibration indicated that the calibration was not over-fitted (Table 5.36 Factor=1 F-value=26.1; Factor=15 F-value=50.4), and the size of the F-value which was an improvement on the individual seasons, indicated that the calibration was more capable of making reliable predictions [8]. This was reinforced by the trend seen in error of cross validation (Table 5.36 Factor=1 SECV=25.37; Factor=15 SECV=21.34), as more factors were added to the calibration there is an improvement in SECV which indicates that prediction is improving as factors were added [9]. The coefficient of determination of the validation set achieved was 0.798 with an uncorrected SEP of 16.92 and Bias of -1.14 (Table 5.36).

#### **5.4.5 Comparison to cited literature**

Work by McGlone *et al* [21,22] highlighted the difficulties of relating spectroscopic data of whole samples to titratable acidity values. In these studies, poorly performing calibrations of titratable acidity were developed, which proved incapable of accurately predicting titratable acidity levels in whole fruit beyond coarse screening (“Satsuma” Mandarins  $R^2=0.65$  [21]; “Royal Gala” apples  $R^2=0.38$ ) [22]. These calibrations achieved coefficient of determination of no better than 0.65, in this case with whole mandarins [21], which was achieved by an indirect association of chlorophyll levels within the skin of the fruit being studied. This is contrasted to the final PLS calibration of processed sultanas, which achieved a coefficient of determination superior to both these examples (Table 5.36 PLS  $R^2=0.88$ ). When more uniform media are examined with NIRS, successful calibrations for titratable acidity can be developed, as can be seen by the fermented milk product work by Navratil *et*

*a/* [23]. In this case yoghurts and filmjолks were analysed using NIRS and calibrations from this data were developed, achieving coefficient of determination of 0.995 using 5 PLS factors. This coefficient of determination was a much stronger prediction of titratable acidity than the calibration for processed sultanas (Table 5.36 PLS  $R^2=0.88$ ), which was largely due to its lack of homogeneity compared with the fermented milk products.

## 5.5 Total lipids

### 5.5.1 First season Total Lipids

Table 5.37 Laboratory parameters summary of season one Lipids

# Samples (in calibration)	# Samples in validation	Range	Mean	Standard Deviation
76(61)	15	0.928-1.654	1.335	0.156

#### 5.5.1.1 MLR Calibration

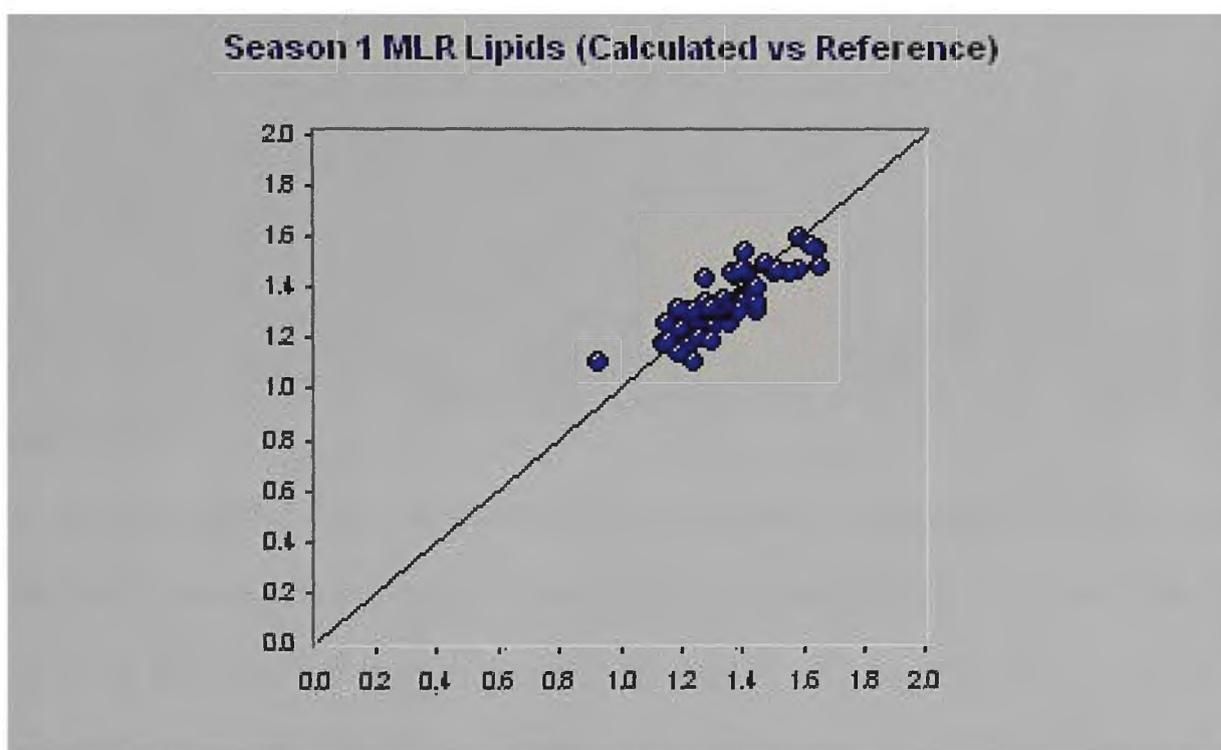


Figure 5.37 A plot of the calibration for percent lipids using MLR

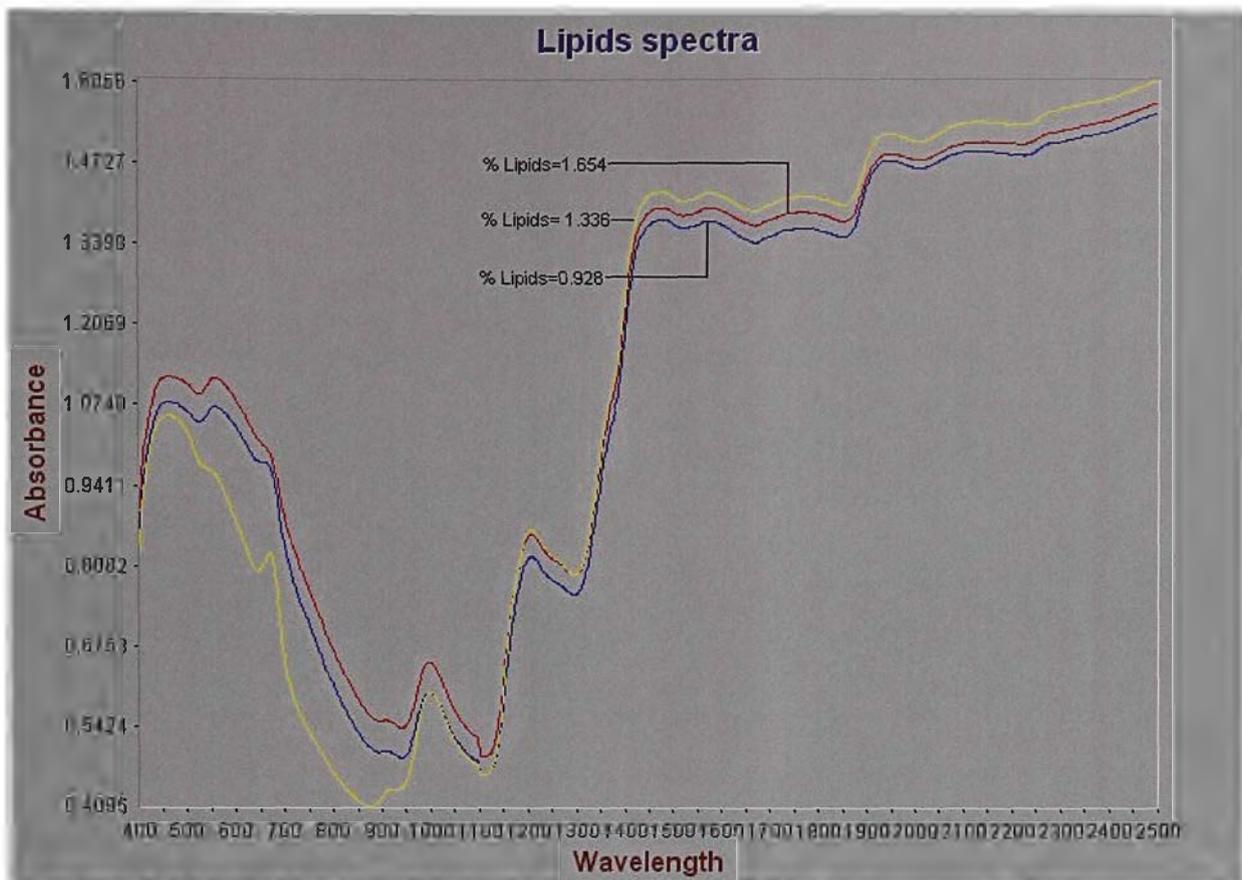


Figure 5.38 Spectra of the minimum (blue), maximum (red) and mean (yellow) samples of percent lipids

A diagram containing three spectra of the titratable acidity calibration is shown above (Figure 5.38) the spectra displayed are of the spectra of the sample with the highest percent lipids value (1.654) in red, the spectra of the sample closest to the mean (1.336) in yellow and the spectra of the sample with the minimum percent lipids value (0.926) in blue. Below is displayed a close up of the region of the first wavelength of the MLR calibration (Figure 5.39).

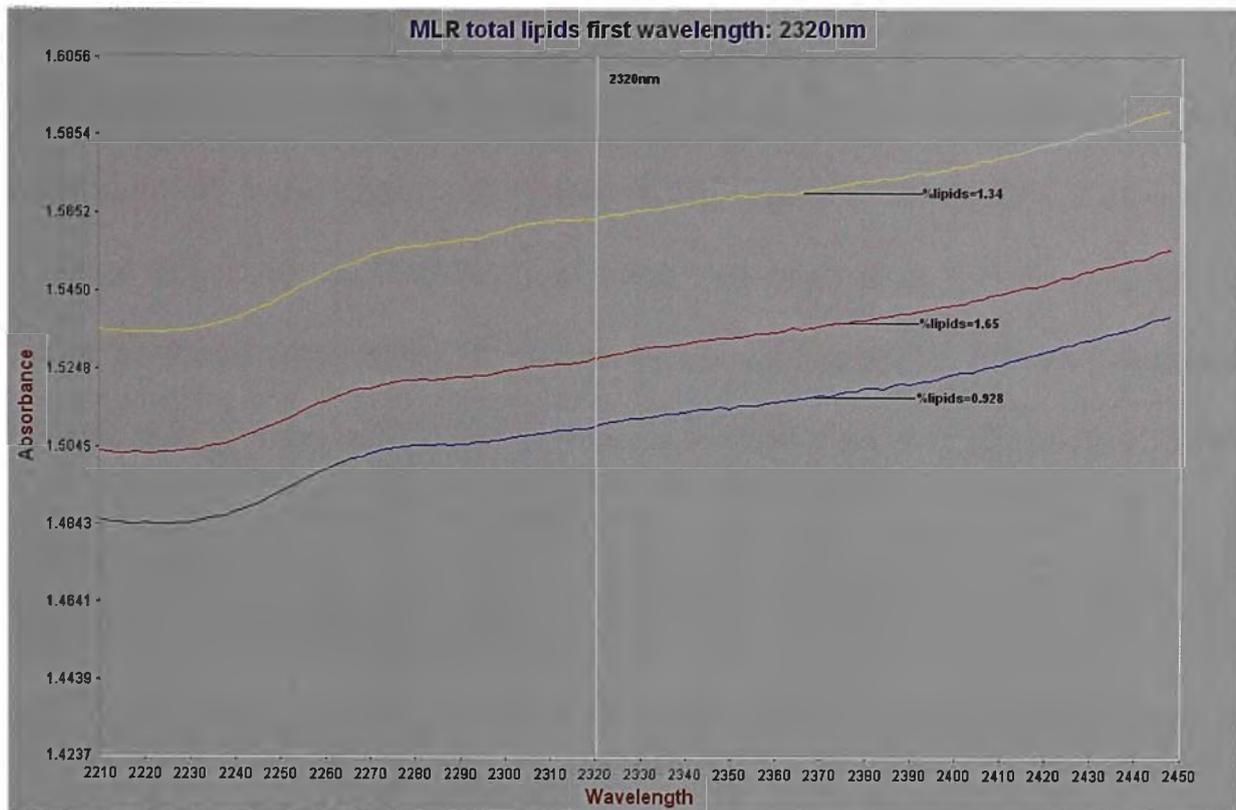


Figure 5.39 Close up of NIR spectra at the first wavelength of the combined MLR calibration

The calibration that was developed for total lipids had summation terms that showed a link to the stretching and deformation wavelengths of C-H bonds [5] of methylene groups and tertiary carbons from the third overtone region [6] (Table 5.38 766nm), a wavelength attributable to C-H vibrations of alkyl groups [6] (2320nm) and a wavelength attributable to secondary and tertiary carbon C-H bonds from the combinations band [6] (2412nm) and a wavelength near the primary amines region [5] (1564nm).

Table 5.38 Wavelengths contributing to the first seasons MLR percent lipids calibration (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

MLR					
Term	Outliers	Wavelength (nm)	R <sup>2</sup> (Cal)	SEC	F-value
1	6	2320	0.335	0.12	25.2
2	“	1564	0.470	0.11	21.8
3	“	766	0.626	0.092	26.8
4	6	2412	<b>0.68</b>	<b>0.086</b>	<b>25.0</b>

A plot of the calibration set is provided above (Figure 5.37). The F-value of the MLR calibration was small, but steadily improved as more terms were added to the calibration, showed that the calibration was still viable (Table

5.38 Term=1 F-value=25.2; Term=4 F-value=25.0) [8]. The resultant coefficient of determination was low, but as more terms were added a steady improvement was observed (Table 5.38 Term=1  $R^2=0.335$ ,  $SEC=0.12$ ; Term=4  $R^2=0.68$ ,  $SEC=0.09$ ), indicating this calibration is not of the same value as those with higher  $R^2$  values, but is applicable for coarse prediction and in combination with other calibrations added another dimension to the assessment even though it is not strong.

### 5.5.1.2 PLS Calibration

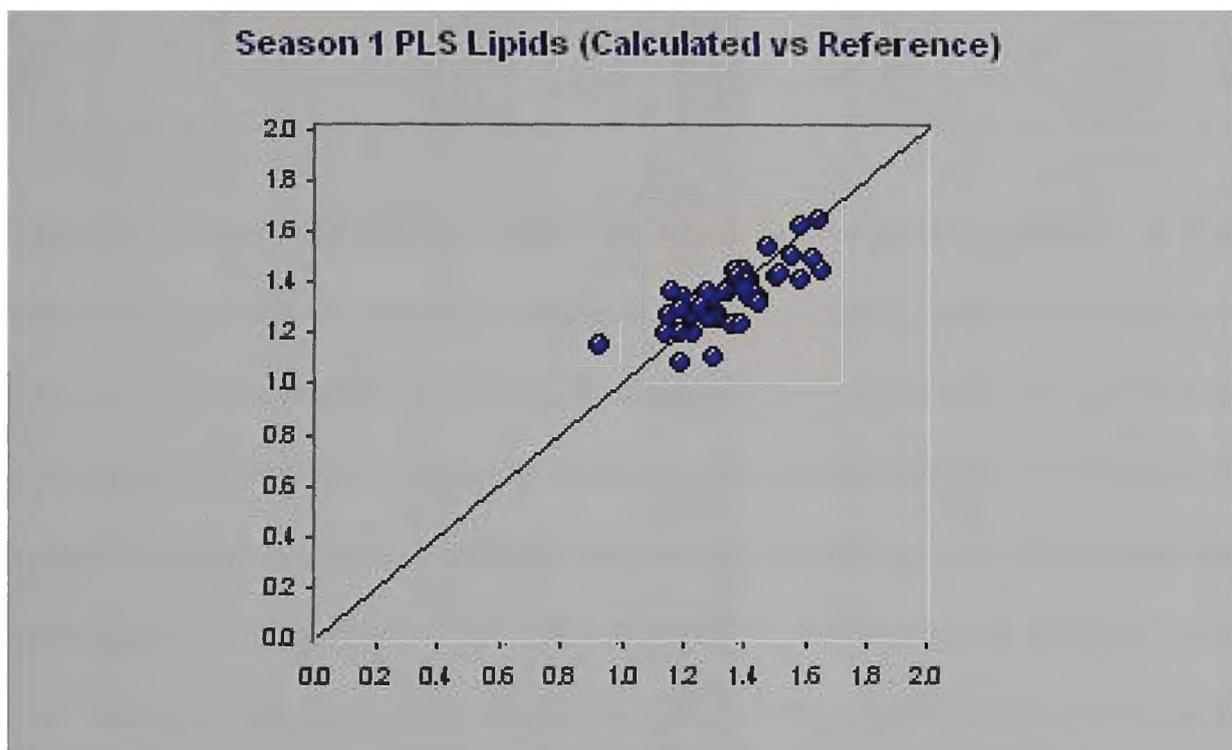


Figure 5.40 A plot of the calibration for percent lipids using PLS

Above is provided a plot of the calibration set of the PLS calibration (Figure 5.40). The contrasting performance of both calibrations shows that MLR has handled this relatively small calibration better than PLS, however both calibrations are sub-standard, their F-values (Table 5.38 MLR F-value=25.0; Table 5.39 PLS F-value=10.6) were very low which indicated a poor reliability of prediction [8], and any further work on this parameter was halted owing to the factor that there was insufficient variation in the lipid content in the

samples. Also their coefficients of determination were poor, MLR showed some usefulness as a coarse screening calibration (Table 5.38  $R^2=0.68$ ,  $SEC=0.09$ ), however PLS did not achieve any level of usefulness but improved as more factors were included into the calibration (Table 5.39 Factor=1  $R^2=0.034$ ,  $SEC=0.144$ ; Factor=6  $R^2=0.59$ ,  $SEC=0.10$ ).

Table 5.39 Factors contributing to the first seasons' PLS percent lipids calibration (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

PLS					
Factors	Outliers	$R^2$ (Cal)	SEC	SECV	F-value
1	6	0.034	0.144	0.160	1.78
2	"	0.149	0.137	0.154	4.29
3	"	0.260	0.129	0.159	5.62
4	"	0.430	0.114	0.166	8.86
5	"	0.538	0.104	0.164	10.70
<b>6</b>	<b>6</b>	<b>0.586</b>	<b>0.100</b>	<b>0.178</b>	<b>10.60</b>

The trend seen with the F-values as more factors were included to the calibration show that the calibration was not over-fitted (Table 5.39 Factor=1 F-value=1.78; Factor=6 F-value=10.6) however the modest size of the F-value also showed that it was incapable of making reliable predictions [8]. The trend seen in the error of cross validation indicates a gradual loss of robustness as more factors were added to the calibration due to an increase in SECV (Table 5.39 Factor=1 SECV=0.16; Factor=6 SECV=0.178), which supported the lack of reliability and sample variation indicated by the low F-value [9].

### 5.5.2 Comparison to cited literature

Work by Gonzales-Martin et al [24] has developed a series of calibrations used to predict total saturated and unsaturated fats and specific fatty acids in Iberian breed swine, using NIR diffuse reflectance spectra of intact subcutaneous pork fat samples using a remote fibre optic probe and spectra

of solvent extracted fat measured by NIRS held in cam-lock cups. The calibrations of total polyunsaturated fats achieved a coefficient of determination of 0.948 with an external validation  $R^2$  of 0.88 for whole samples, the mono-unsaturated showed a  $R^2$  of 0.897 and corresponding validation of 0.74, while the saturated fat calibration  $R^2$  of 0.958 and an external validation of 0.9.

The solvent extracted sample calibration for total polyunsaturated fats gave a  $R^2$  of 0.958 and an external validation of 0.85, total mono- gave a coefficient of determination 0.939 with an external validation of 0.9 and total saturated fats achieved a coefficient of determination of 0.986 and an external validation of 0.97 [24]. Surprisingly the results obtained for intact subcutaneous pork fat samples performed better than processed sultanas (Table 5.38 MLR  $R^2=0.68$ ; Table 5.39 PLS  $R^2=0.59$ ), as it would be expected to be less homogeneous, however the variation in fat types in samples would improve the chances of achieving a valid predictive model.

While these calibrations are not total lipid calibrations, they are analytes of similar chemical nature to those studied in this project and therefore adsorb at similar wavelengths to those of the calibrations that were developed for whole processed sultanas.

## 5.6 Initial Conclusions

A series of samples from three seasons from the processors in the Sunraysia Region were collected, of as wide a variety of fruit crown grades and maturities as were available during the seasons of the project.

The samples collected were analysed with Kjeldahl protein, water activity, percent lipids and titratable acidity. These constituent values were then used to develop calibrations using both MLR and PLS regression analysis techniques combined with NIRS diffuse reflectance spectroscopy. The MLR calibration developed for water activity by dew point achieved a coefficient of determination of 0.936 and standard error of calibration of 0.0059 and a coefficient of determination for validation of 0.965. The PLS calibration of the combined seasons results achieved a coefficient of determination of 0.949 with an standard error of 0.0053 and achieved a prediction coefficient of determination of 0.957. The titratable acidity calibration developed with MLR yielded a coefficient of determination of 0.736 with the corresponding standard error of calibration of 14.5, and gave a prediction coefficient of determination of 0.70. The PLS calibration for titratable acidity achieved a coefficient of determination of 0.877 and a standard error of calibration of 10.33 and the resultant prediction coefficient of determination of 0.798. Kjeldahl protein gave a MLR calibration with a coefficient of determination of 0.793 and standard error of calibration of 0.18 and a validation coefficient of determination of 0.759. The PLS calibration developed for Kjeldahl protein achieved a coefficient of determination of 0.859 with a standard error of calibration, and gave a prediction coefficient of determination of 0.764.

Developing a successful calibration for percent lipids for processed sultanas was hindered by the fact that the processed sultanas have a finishing oil added which confers a certain degree of uniformity. The lack of a viable range of sultanas with high and low lipid content resulted in poor results. Coefficients of determination for MLR and PLS were 0.68 and 0.586 respectively. Lipid analysis was thus not continued after the first season.

The rapid assessment of processed samples using NIRS on random batches of samples will begin as soon as the industry is ready to adopt the technique, a limited trial of the calibrations will occur in the coming season.

Addressing the experimental hypothesis:

- The resultant calibrations of this chapter have shown that rapid, objective assessment of sultana quality parameters is feasible, and would aid in the rapid analysis of parameters important to meet future processed food labelling regulations.

## 5.7 Bibliography

1. Frank, D, Gould, I, Millikan, M. "Browning Reactions During Storage of Low Moisture Australian Sultanas: Further Evidence for Arginine Mediated Maillard Reactions During Storage and Some Effects of Vine Shading and Harvest Date", *Australian Journal of Grape and Wine Research* 2004;10(3):182-195.
2. Frank, D, Gould, I, Millikan, M. "Browning Reactions During Storage of Low Moisture Australian Sultanas: Evidence for Arginine Mediated Maillard Reactions", *Australian Journal of Grape and Wine Research* 2004;10(2):151-163.
3. Kochhar, SP Rossell, RB. "A Vegetable Oiling Agent for Dried Fruit", *Journal of Food Technology*. 1982;17:661-668.
4. Uhlig, BA, Clingeffer, PR. "Influence of Grape (*Vitis Vinifera* L.) Berry Maturity on Dried Fruit Colour" *Journal of Horticultural Science and Biotechnology* 1998; 73(3): 329-339
5. The Chart '*Near Infrared Adsorptions: Major analytical bands and their relative peak positions*'. Silver Spring (MD): NIRSystems, Inc; 1992.
6. *Near Infrared technology in the agricultural and food industries*. Second Edition. Edited by P Williams and C Norris. St. Paul (Mi): American Association of Cereal Chemists; 2001;p.33-36
7. Osborne, B, Fearn, T, Hindle, P. *Near Infrared Spectroscopy with applications in food and beverage analysis*. 2<sup>nd</sup> Edition. Singapore; Longman Scientific and Technical; 1993.
8. Mark, H. *Principles and practice of spectroscopic calibration*. Brisbane: J Wiley and sons, Inc; 1991; p 31

9. *Near Infrared technology in the agricultural and food industries*. Second Edition. Edited by P Williams and C Norris. St. Paul (Mi): American Association of Cereal Chemists; 2001;p.154.
10. Beebe, KR, Pell, RJ, Seasholtz, MB. *Chemometrics, a practical guide*. Brisbane: J Wiley and Sons, Inc.;1998; p.335-336.
11. Huxoll, CC. "Near infrared analysis potential for measuring the water activity of raisins". *Conference Proceedings of the IFT Annual Meeting*. 1995: p70.
12. Huxoll, CC. "Assessment of Near Infrared (NIR) Diffuse Reflectance analysis for measuring Moisture and Water Activity in Raisins", *Journal of Food Processing and Preservation* 2000;24(4): 315-333.
13. Flinn, PC, Black, RG, Iyer, L, Brouwer, JB, Meares, C. "Estimating the Food Processing Characteristics of Pulses by Near Infrared Spectroscopy, using Ground or Whole Samples", *Journal of Near Infrared Spectroscopy* 1998;6:213-220.
14. Blakeney, AB, Welsh, LA, Sharman, JP, Ronalds, JA, Reece, JE. "Rice Moisture Analysis by NIT". In: Batten, GD, Flinn, PC, Welsh, LA, Blakeney, AB, editors. *Leaping ahead with Near Infrared Spectroscopy*. Melbourne: NIRSG, RACI; 1995; p.316-323.
15. Mark, H. *Principles and practice of spectroscopic calibration*. Brisbane: J Wiley and sons, Inc; 1991;p92-93.
16. Fassio, A, Cozzolino, D. "Non-destructive Prediction of Chemical Composition in Sunflower Seeds by Near Infrared Spectroscopy", *Industrial Crops and Products* 2004;20(3):321-329

17. Confalonieri, M, Lombardi, G, Bassignana, M, Odoardi, M. "Analysis of Quality Constituents of Natural Alpine Swards with Near Infrared Reflectance Spectroscopy", *Journal of Near Infrared Spectroscopy* 2004;12(6):411-417.
18. Kays, SE, Barton, FE, Windham, WR. Predicting Protein Content by Near Infrared Reflective Spectroscopy in Diverse Cereal Food Products", *Journal of Near Infrared Spectroscopy* 2000;8(1):35-43.
19. White, M, Rouvinen-Watt, K. "Near Infrared Evaluation of Wet Mink Diets", *Animal Feed Science and Technology*. 2004;111(1-4):239-246.
20. Kasimatis, AN, Vilas, EP Jr., Swanson, FH, Baranek, PB. "Relationship of Soluble Solids and Berry Weight to Airstream Grades", *American Journal of Enology and Viticulture* 1977; 28(1): 8-15.
21. McGlone, VA, Fraser, DG, Jordan, JB, Kunnemeyer, R. "Internal Quality Assessment of Mandarin Fruit by Visible-Near Infrared Spectroscopy", *Journal of Near Infrared Spectroscopy* 2003;11(5):323-332.
22. McGlone, VA, Jordan, RB, Martinsen, PJ. "Visible-NIR Estimation at Harvest of Pre- and Post-storage Quality Indices for Royal Gala Apple", *Post-harvest Biology and Technology* 2002;25(2):135-144.
23. Navratil, M, Cimander, C, Mandenius, CF. "Online Multi-Sensor Monitoring of Yoghurt and Filmjolk Ferments on Production Scale", *Journal of Agricultural and Food Chemistry* 2004;52(3):415-420.
24. Gonzales-Martin, I, Gonzales-Perez, C. Hernandez-Mendez, J, Alvarez-Garcia, N. "Determination of Fatty Acids in the Subcutaneous

Fat of Iberian Breed Swine by Near Infrared Spectroscopy (NIRS) With a Fibre Optic Probe”, *Meat Science* 2003;65(2):713-719.

## **6.0 Results and discussion (III): Quality Assurance of Unprocessed Sultanas - Airstream Sorting**

### **6.1 Introduction**

Previously in this thesis, methods used to develop a calibration to measure berry maturity using Near Infrared Spectroscopy have already been discussed in Chapter 5. In the current chapter the results obtained from the airstream sorting study together with the observations drawn from these results will be presented, the ultimate outcome being the implementation of this calibration as a quality parameter of unprocessed fruit.

While there is a range of quality parameters that will be analysed in this work, it is expected that the most significant will be titratable acidity, as this has a strong proven relationship with maturity. [1,2,3,4,5]

Experimental hypothesis:

- A series working calibrations associated with berry maturity by airstream sorting, that would enable the use of CIE values as well as water activity to choose an optimal time for the harvesting of fruit, not only maturity, but a time chosen that ensures a mature product that is light in colour and full bodied.

## 6.2 Preliminary interpretation of data

### 6.2.1 Airstream sorting: trends within parameters

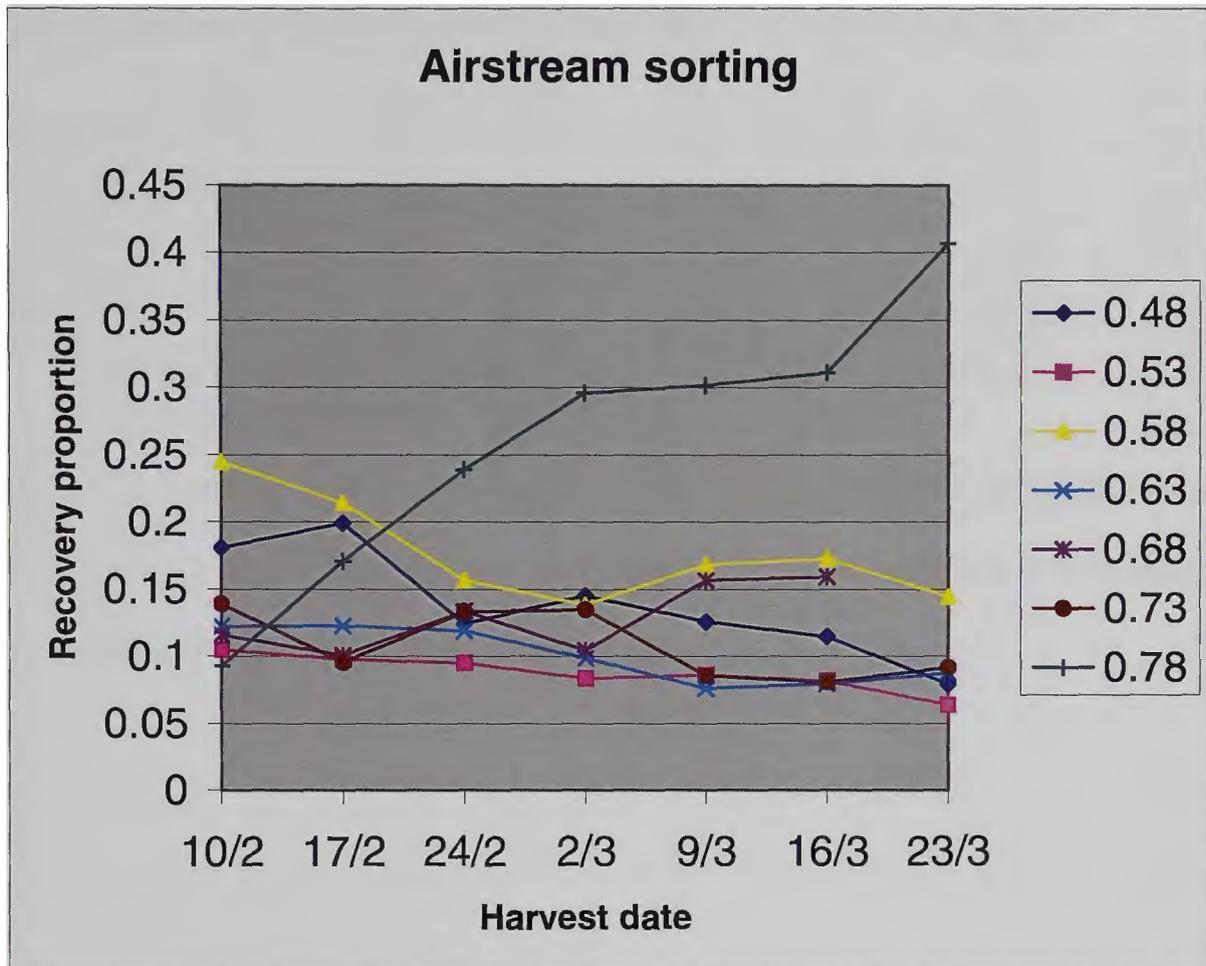


Figure 6.1 a graph of fruit proportions recovered from samples harvested at different dates.

Comparing the mass recovered after airstream sorting (AS) to the airflow pressure (measured in inches of water, analogous to millimetres of mercury) (Table 6.1, Figure 6.1), related to maturity, compared to date, there was the expected trend of an overall decrease in the proportion of lower maturity fruit over time, and conversely there was an overall increase in the proportion of higher maturity fruit over the same time interval.

Unexpectedly, at an airflow pressure of 0.68 inches of water, there was a general increase in proportion of mature fruit over time while at both 0.73 and 0.63 inches of water there was a general decrease in the amount of sample

collected. The 0.73 inches of water value is in itself unusual, as it is a relatively high maturity fraction but the mass recovered at this setting decreased.

Table 6.1 Recovery proportion values of fruit at pressure settings for harvest dates.

Date	0.48" water	0.53" water	0.58" water	0.63" water	0.68" water	0.73" water	0.78" water
10/02/2004	0.181	0.105	0.245	0.122	0.116	0.139	0.093
17/02/2004	0.199	0.098	0.214	0.122	0.101	0.095	0.171
24/02/2004	0.124	0.095	0.157	0.119	0.133	0.133	0.239
2/03/2004	0.145	0.083	0.138	0.099	0.104	0.135	0.296
9/03/2004	0.126	0.086	0.168	0.076	0.157	0.085	0.302
16/03/2004	0.114	0.082	0.173	0.080	0.159	0.081	0.311
23/03/2004	0.079	0.064	0.144	0.087	- (*)	0.092	0.407

(\*)Note, a recovered fraction of fruit of a air pressure setting of 0.68" of water from the fruit harvested on the 23<sup>rd</sup> of March was not available.

While this study is limited in its scope, it did show some interesting behaviour of fruit during maturation that certainly did not fit commonly held expectations of fruit during this time, for more detail see below.

## 6.2.2

### Proportion of fruit versus Harvest time

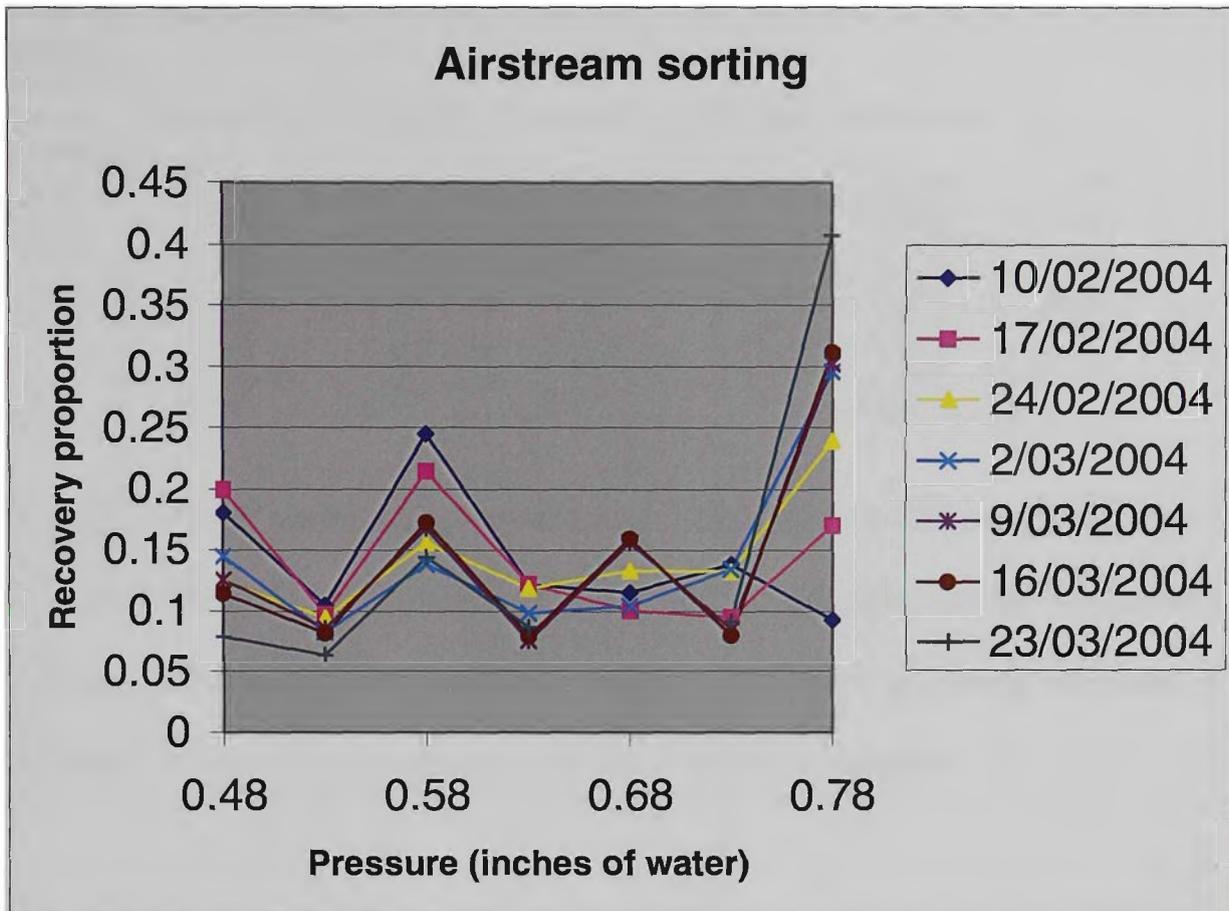


Figure 6.2 fruit fractions recovered at an AS pressure setting of fruit harvested at different dates

In Section 6.2.1, it was observed that many of the trends seen in the proportion of fruit recovered at different harvest dates gave some results that were unexpected, this was supported if the data were examined by comparing the change of fruit recovered at airstream sorting settings, compared to the harvest time (Table 6.2 Figure 6.2).

At 0.48 and 0.53 inches of water, there was a trend of general decrease in recovered mass of fruit over time, while at 0.58 inches of water there was initially a strong decrease followed by a period of increase in recovered mass of fruit that plateaued then decreases, while at a pressure of 0.63 inches of water there was a general decrease in recovered mass of fruit over time. At 0.68 inches of water, there was a period of flux followed by a brief period of increased proportion of fruit recovered proceeded by a decline.

Table 6.2 Recovered fruit fractions compared to date of harvest versus airstream sorting settings

Pressure (inches of water)	10/02/2004	17/02/2004	24/02/2004	2/03/2004	9/03/2004	16/03/2004	23/03/2004
0.48	0.181	0.199	0.124	0.145	0.126	0.114	0.079
0.53	0.105	0.098	0.095	0.083	0.086	0.082	0.064
0.58	0.245	0.214	0.157	0.138	0.168	0.173	0.144
0.63	0.122	0.122	0.119	0.099	0.076	0.080	0.087
0.68	0.116	0.101	0.133	0.104	0.157	0.159	-(*)
0.73	0.139	0.095	0.133	0.135	0.085	0.081	0.092
0.78	0.093	0.171	0.239	0.296	0.302	0.311	0.407

At 0.73 inches of water, initially there was a decrease in fruit present followed by a brief increase that preceded a decline in fruit collected at that airflow pressure. Whereas 0.78 inches of water went through a steady increase in fruit that briefly reached a plateau then continued to increase.

### 6.2.3 Colour versus Harvest time

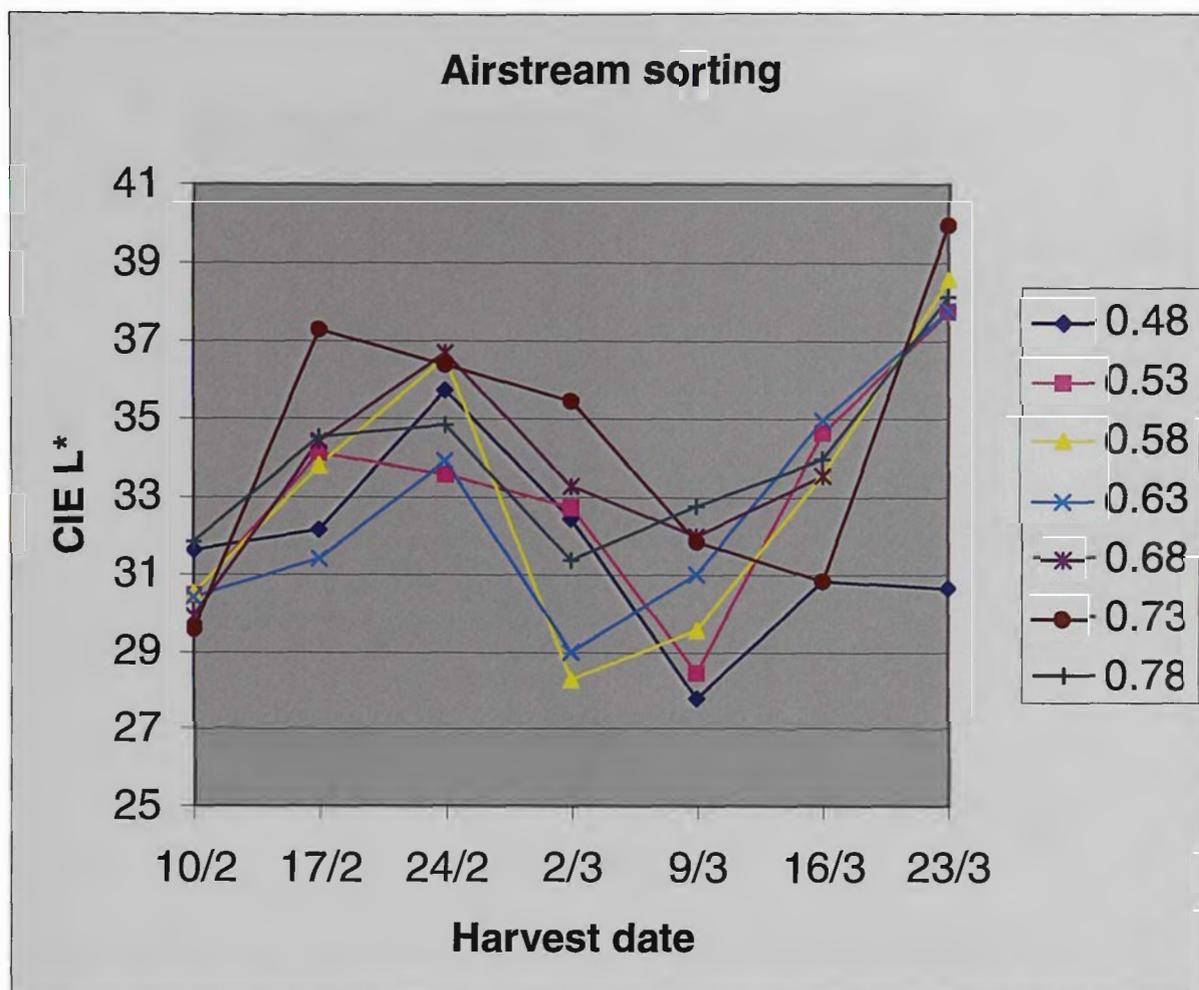


Figure 6.3 harvest dates versus average CIE L\* values at airstream settings

As can be seen in the data presented above (Table 6.3, Figure 6.3), initially, an increase in L\* was observed across all AS fractions, this increase reached a plateau and decreased and then continued to increase, except in 0.40 the least mature fraction that reached a plateau after the second increase.

Table 6.3 AS setting at harvest date versus average CIE L\* value

Date	0.48" water	0.53" water	0.58" water	0.63" water	0.68" water	0.73" water	0.78" water
10/02/2004	31.62	30.47	30.53	30.42	29.92	29.6	31.84
17/02/2004	32.15	34.11	33.8	31.41	34.42	37.3	34.53
24/02/2004	35.73	33.57	36.59	33.91	36.7	36.38	34.84
2/03/2004	32.44	32.74	28.31	29.02	33.29	35.46	31.38
9/03/2004	27.81	28.47	29.58	31.01	32	31.84	32.76
16/03/2004	30.83	34.64	33.58	34.97	33.55	30.83	33.97
23/03/2004	30.64	37.76	38.56	37.79	- (*)	39.97	38.16

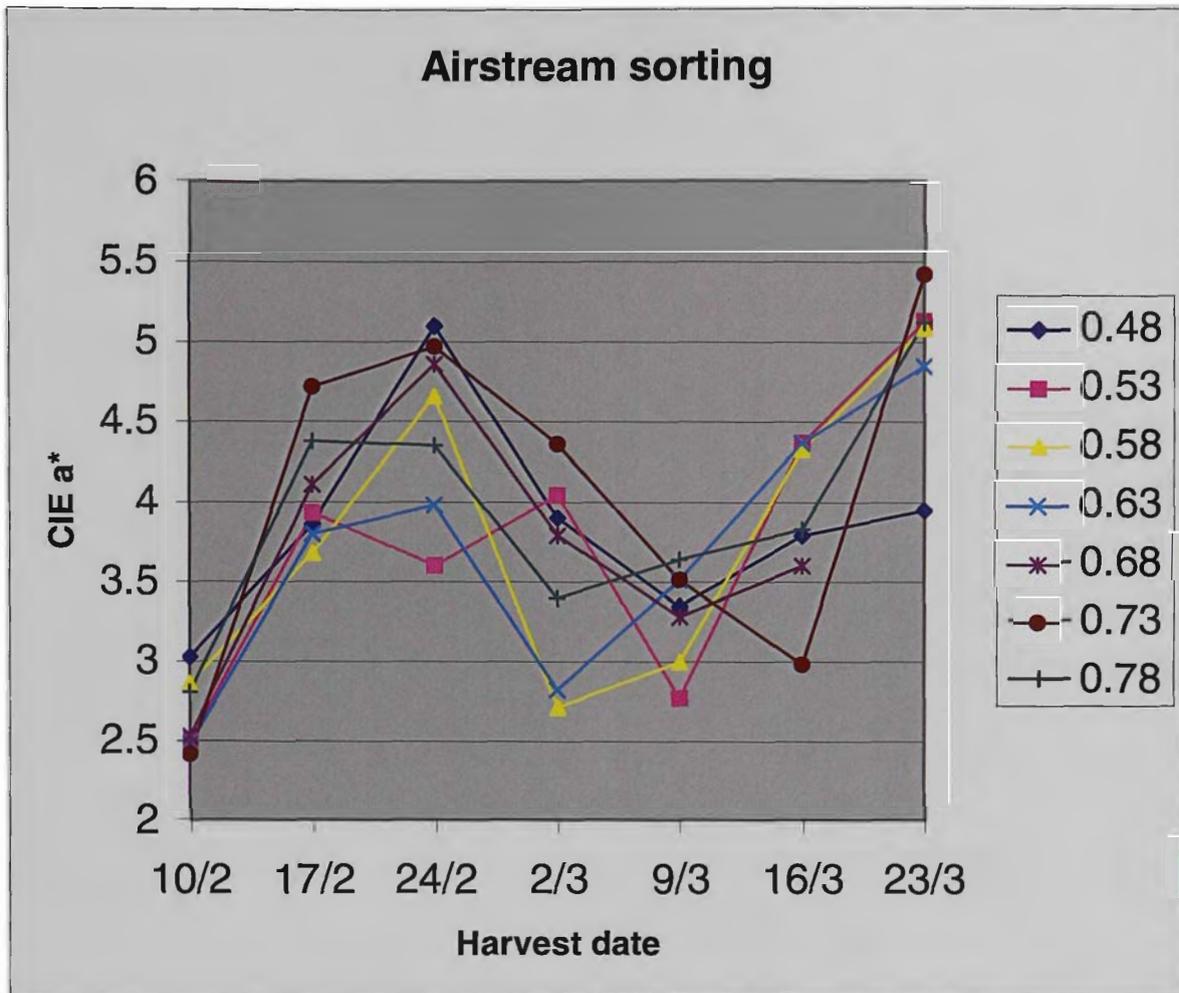


Figure 6.4 harvest dates versus average CIE a\* values at airstream settings

In the case of a\* (Table 6.4, Figure 6.4), it followed a very similar trend to L\*, in that over all fractions there was an initial increase then a decline and finally an increase. In most cases this final increase in L\* was considerable while for the least mature fraction, the increase in L\* was more gradual.

Table 6.4 AS setting at harvest date versus average CIE a\* value

Date	0.48" water	0.53" water	0.58" water	0.63" water	0.68" water	0.73" water	0.78" water
10/02/2004	3.03	2.5	2.86	2.48	2.53	2.42	2.81
17/02/2004	3.86	3.93	3.68	3.8	4.11	4.72	4.38
24/02/2004	5.1	3.6	4.66	3.98	4.86	4.97	4.35
2/03/2004	3.9	4.04	2.71	2.82	3.79	4.36	3.39
9/03/2004	3.35	2.77	3	3.51	3.28	3.51	3.64
16/03/2004	3.79	4.37	4.33	4.37	3.6	2.98	3.83
23/03/2004	3.95	5.13	5.09	4.85	- (*)	5.42	5.13

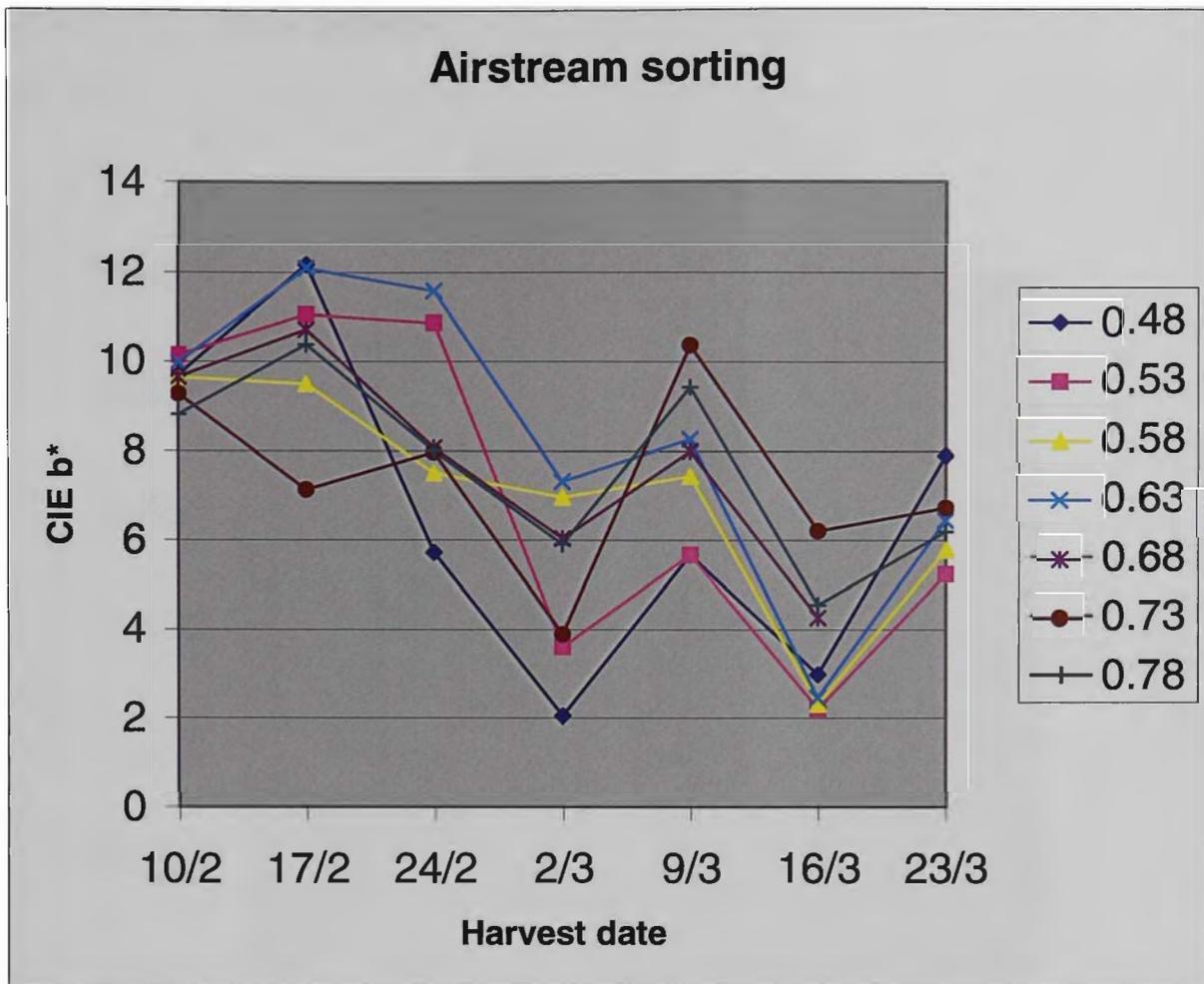


Figure 6.5 Harvest dates versus average CIE b\* values at airstream settings

The case of b\* (Table 6.5, Figure 6.5) followed a generally downward trend, which briefly undergoes a sharp increase at week 5, followed by an equally sharp decline, finally followed by a more gradual increase.

Table 6.5 AS setting at harvest date versus average CIE b\* value

Date	0.48" water	0.53" water	0.58" water	0.63" water	0.68" water	0.73" water	0.78" water
10/02/2004	9.7	10.14	9.65	9.98	9.66	9.28	8.82
17/02/2004	12.17	11.05	9.5	12.09	10.71	7.13	10.38
24/02/2004	5.72	10.86	7.5	11.59	8.06	7.96	7.99
2/03/2004	2.06	3.6	6.97	7.32	6.04	3.88	5.90
9/03/2004	5.68	5.68	7.45	8.27	7.99	10.37	9.43
16/03/2004	2.99	2.21	2.32	2.47	4.27	6.21	4.55
23/03/2004	7.9	5.24	5.79	6.45	- (*)	6.73	6.20

6.2.4 Water activity

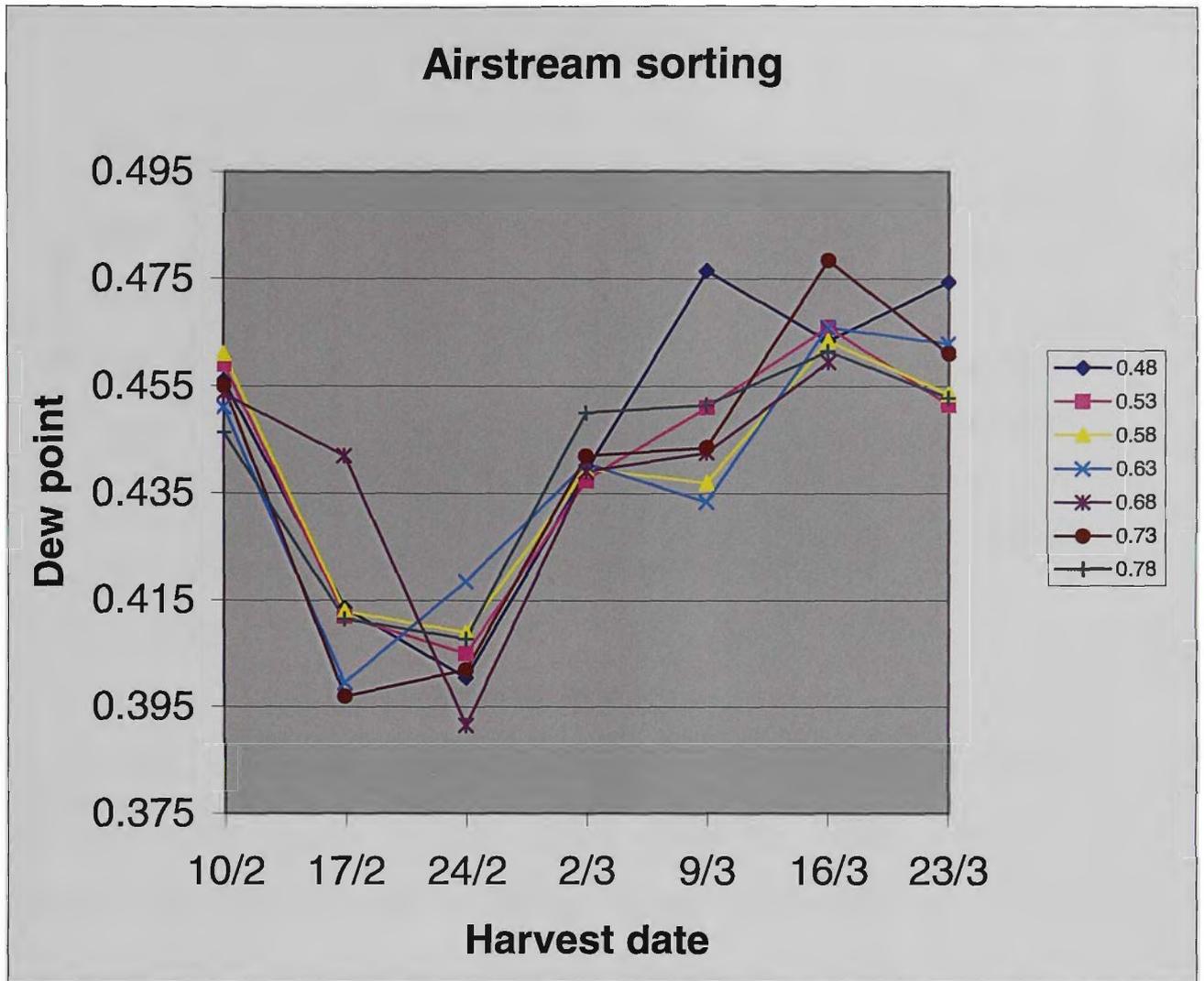


Figure 6.6 harvest dates versus dew point (water activity) values at airstream settings

As can be seen in the data presented above (Table 6.6 Figure 6.6) the amount of unbound water present in all samples decreased sharply. This sharp decrease was followed by a more gradual and sustained increase over time.

Table 6.6 Harvest dates versus average dew point (water activity) values at airstream settings

Date	0.48" water	0.53" water	0.58" water	0.63" water	0.68" water	0.73" water	0.78" water
10/02/2004	0.456	0.459	0.461	0.451	0.454	0.455	0.446
17/02/2004	0.414	0.412	0.413	0.400	0.442	0.397	0.412
24/02/2004	0.401	0.405	0.409	0.419	0.392	0.402	0.408
2/03/2004	0.439	0.438	0.440	0.441	0.439	0.442	0.450
9/03/2004	0.477	0.451	0.437	0.434	0.443	0.444	0.452
16/03/2004	0.464	0.466	0.464	0.466	0.460	0.479	0.462
23/03/2004	0.475	0.452	0.454	0.463	- (*)	0.461	0.453

## 6.2.5 Titratable acidity

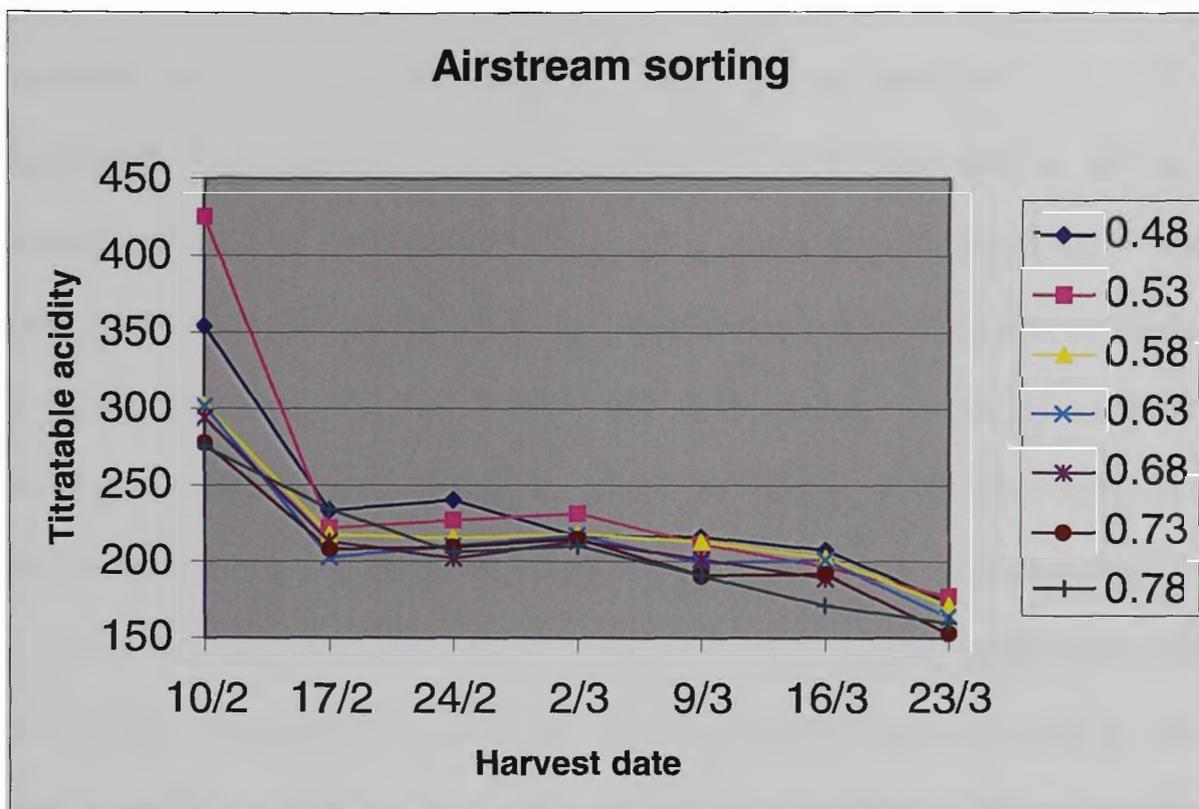


Figure 6.7 harvest dates versus titratable acidity (mL of 0.1M NaOH/100g sample) values at airstream settings

The pattern followed for titratable acidity (Table 6.7 Figure 6.7) over the course of the study was one of decline; initially rapidly which tailed off to a more gradual decrease for the remainder of the study.

Table 6.7 Harvest dates versus average TA values at airstream settings

Date	0.48" water	0.53" water	0.58" water	0.63" water	0.68" water	0.73" water	0.78" water
10/02/2004	353.9	425.6	302.1	301.8	294.0	277.6	275.7
17/02/2004	233.4	221.7	217.5	202.7	213.1	208.0	234.3
24/02/2004	240.4	226.9	215.3	210.1	202.0	209.3	206.0
2/03/2004	216.6	231.3	217.8	217.0	213.5	214.6	210.1
9/03/2004	216.1	211.7	212.9	199.0	201.4	190.7	190.9
16/03/2004	207.4	196.3	202.5	200.7	188.7	191.9	171.4
23/03/2004	173.6	177.1	171.2	164.1	- (*)	153.0	159.9

### **6.2.6 Titratable acidity calibration development**

Using the initial batch of airstream-sorted samples combined with samples unsorted by the airstream sorter, a calibration of titratable acidity was developed. This calibration used a total of 134 samples with a standard deviation of 44.9 ml 0.1M NaOH/100g and a range of samples, which varied from 153 to 467 ml 0.1M NaOH/100g. Calibrations were developed by means of Vision Software Version 2.22 (Foss NIRSystems, propriety software, Denmark) using spectra collected over the range of 400-2500nm. Pre-treatments for the calibrations in all cases were N-point smooth and 2<sup>nd</sup> derivative. In both PLS and MLR, it was found that little modification was required to fine tune the calibrations, so two outliers were removed in each case. The MLR utilised 6 summation terms and PLS used 10% of the total number of samples in the validation set, to guide how many factors were required to optimise the spectra. In this case, it was decided not to remove redundant samples from the calibration for two reasons, firstly the calibration was intended to be an operational calibration and its size limits the excessive removal of samples [6]. Secondly, on examination of the calibrations histogram, it was seen that the calibration had a diverse range of results with a comparatively even distribution [7].

## 6.3 Airstream-sorting NIRS calibration development

### 6.3.1 CIE L\*

Table 6.8 Laboratory parameters summary of CIE L\*

# Samples (in calibration)	# Samples in validation	Range	Mean	Standard Deviation
134(107)	27	27.81-39.97	34.47	2.76

#### 6.3.1.1 MLR calibration of CIE L\*

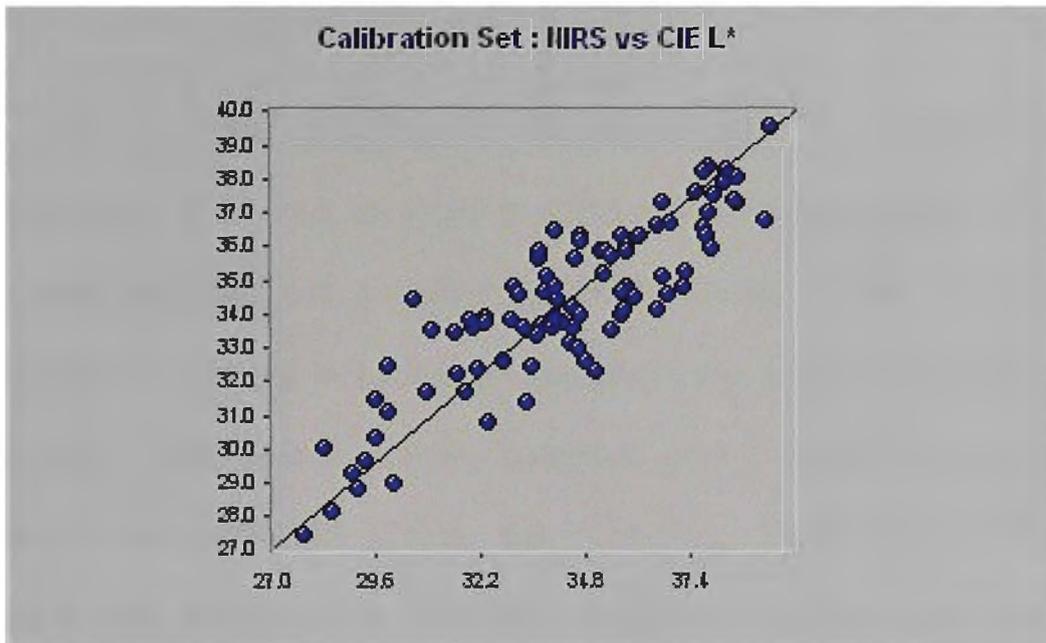


Figure 6.8 A plot of the MLR calibration for CIE L\*

The following assignments are theoretical and as such there are a number of possible origins consistent with the complex matrix of the medium being studied. Further investigation beyond the scope of this project would need to be completed to confirm these assignments. Refer to section 4.6 for further details.

The first and third terms (Table 6.9 496nm, 484nm) of this calibration were found within the visible region of the spectrum, and this constituent was directly related to the colour of the fruit [8]. The other assignments, however, were more difficult to account for, terms two (1154nm) and five (2284nm) were attributed to C-H stretching of methyl groups within organic compounds

in the sample [9,10]. Terms four and six (1834nm, 1552nm) were attributed to O-H stretch in organic hydroxy groups [9,10].

Table 6.9 Wavelengths contributing to the airstream sorting MLR calibration of CIE L\* (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

MLR								
Terms	Outliers	$\lambda$ (nm)	R <sup>2</sup> (Cal)	SEC	F-value	R <sup>2</sup> (Val)	Bias	SEP
1	-	496	0.406	2.19	59.36	-	-	-
2	-	1154	0.519	1.99	46.42	-	-	-
3	-	484	0.619	1.78	46.13	-	-	-
4	-	1834	0.672	1.66	43.01	-	-	-
5	-	2284	0.730	1.52	44.79	-	-	-
<b>6</b>	<b>11</b>	<b>1552</b>	<b>0.765</b>	<b>1.42</b>	<b>44.45</b>	<b>0.778</b>	<b>-1.27</b>	<b>2.18</b>

The calibration of L\* was developed using measurements taken from 134 unprocessed samples using a Minolta CR-300 Chromameter. The range of this calibration assisted in making a comparatively small calibration predict successfully. With careful outlier selection during calibration, a suitable correlation was achieved (Table 6.9 R<sup>2</sup>=0.765, SEC=1.42), a diagram illustrating the calibration is included (Figure 6.8). The calibration also successfully predicted the validation set (Table 6.9 R<sup>2</sup>=0.778). The validation set achieved an uncorrected SEP of 2.18 and a Bias of -1.27 (Table 6.9). Both the coefficient of determination and SEC improved (Table 6.9 Term=1 R<sup>2</sup>=0.406, SEC=2.19; Term=6 R<sup>2</sup>=0.765, SEC=1.42) as terms were added to the calibration. The F-value of the calibration showed a slight decrease in size, however the final value was large indicating the calibration was reliable (Table 6.9 Term=1 F-value=59.36; Term=6 F-value=44.45) [11].

### 6.3.1.2 PLS calibration of CIE L\*

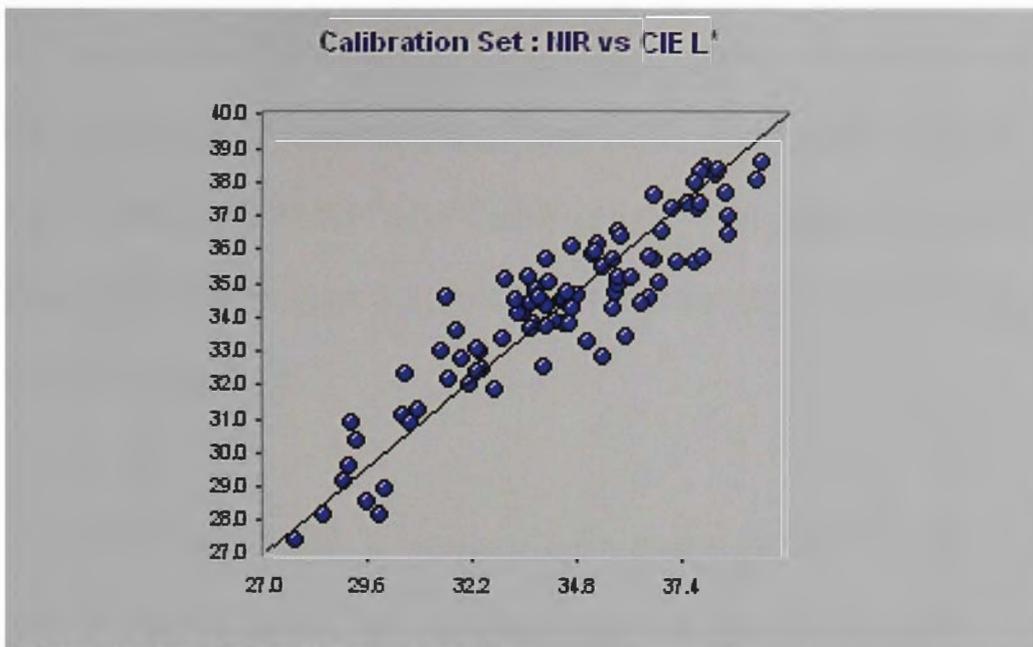


Figure 6.9 A plot of the PLS calibration for CIE L\*

Due to the importance of the CIE coordinate L\* for fruit maturity [5], a calibration of CIE L\* using PLS was developed. As more factors were included into the PLS calibration for CIE L\* steady improvements in both the coefficient of determination and SEC were seen (Table 6.10 F=1 R<sup>2</sup>=0.063, SEC=2.71; F=10 R<sup>2</sup>=0.834, SEC=1.14).

Table 6.10 Factors contributing to the airstream sorting PLS calibration of CIE L\*(Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

PLS								
Factors	Outliers	R <sup>2</sup> (Cal)	SEC	SECV	F-value	R <sup>2</sup> (Val)	Bias	SEP
1	11	0.063	2.71	2.75	5.89	-	-	-
2	"	0.199	2.52	2.64	10.7	-	-	-
3	"	0.356	2.27	2.49	15.6	-	-	-
4	"	0.440	2.12	2.39	16.5	-	-	-
5	"	0.517	1.98	2.36	17.8	-	-	-
6	"	0.594	1.83	2.34	20.0	-	-	-
7	"	0.682	1.63	2.34	24.8	-	-	-
8	"	0.757	1.47	2.34	31.1	-	-	-
9	"	0.805	1.29	2.24	36.3	-	-	-
<b>10</b>	<b>11</b>	<b>0.834</b>	<b>1.14</b>	<b>2.23</b>	<b>39.3</b>	<b>0.812</b>	<b>-1.15</b>	<b>1.99</b>

SECV also showed a steady improvement with the development of the calibration (Table 6.10 F=1 SECV=2.75; F=10 SECV=2.23) [12] while the F-value increased, both are indicative of increased reliability of the calibration

(Table 6.10 F=1 F-value=5.89; F=1 F-value=39.3) [11]. When compared to the MLR calibration of this data set, the PLS showed clear improvements in both the coefficient of determination and SEC (Table 6.9 MLR  $R^2=0.765$  SEC=1.42; Table 7.10 PLS  $R^2=0.83$  SEC=1.14). The validation coefficient of determination achieved was 0.812 with an uncorrected SEP of 1.99 and a Bias of -1.15 (Table 6.10).

Flinn *et al* [13] used NIRS to develop calibrations to assess the quality of chickpeas and ground peas for a series of parameters including CIE  $L^*$   $a^*$  and  $b^*$ , and obtained correlation coefficients for whole field and chickpeas of 0.97 and 0.96 for  $L^*$ . This result is higher when compared to the calibration developed from unprocessed sultanas from the maturity study (Table 6.10 MLR  $R^2=0.765$ ; Table 7.10 PLS  $R^2=0.834$ ). In the study by Flinn *et al*, however, the range of samples used in calibration development was much larger, and PLS was used instead of MLR. It should be noted that whole and ground chickpeas are a more homogeneous product than sultanas so a higher correlation coefficient would be expected [13].

### 6.3.2 CIE a\*

Table 6.11 Laboratory parameters summary of CIE a\*

# Samples (in calibration)	# Samples in validation	Range	Mean	Standard Deviation
134(107)	27	2.36-5.94	4.13	0.7673

#### 6.3.2.1 MLR Calibration of CIE a\*

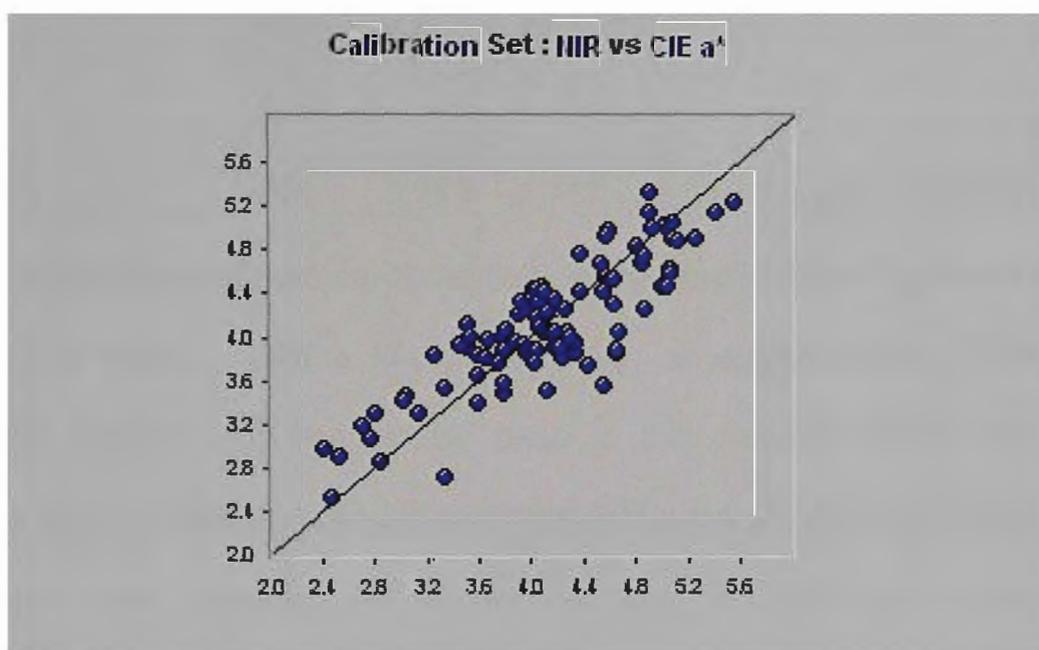


Figure 6.10 A plot of the MLR calibration for CIE a\*

Terms within the visible region were again strongly represented with terms two, five and six present (Table 6.12 486nm, 570nm and 496nm) in the visible region of the spectra [8]. The most significant term (786nm), however, was from near the visible region and was attributed to the stretching vibration [8] of an aromatic amine N-H bond [9]. The two remaining frequencies three and four (1822nm and 1856nm) were associated with hydroxy groups O-H bond stretching vibrations [8].

The standard deviation and range of a\* (Table 6.11 SD=0.767, Range=2.35-5.94) was comparable to larger calibrations of its type. Typically a\* was found to be one of the more challenging constituents because typically the

comparative range of colour values was relatively small while the analytical error remained the same.

Table 6.12 Terms contributing to the airstream sorting MLR calibration of CIE a\* (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

MLR								
No. Terms	Outliers	$\lambda$ (nm)	R <sup>2</sup> Cal	SEC	F-value	R <sup>2</sup> (Val)	Bias	SEP
1	11	786	0.238	0.603	28.75	-	-	-
2	"	486	0.408	0.534	31.41	-	-	-
3	"	1822	0.507	0.49	30.90	-	-	-
4	"	1856	0.612	0.437	35.12	-	-	-
5	"	570	0.653	0.416	33.12	-	-	-
<b>6</b>	<b>11</b>	<b>496</b>	<b>0.717</b>	<b>0.378</b>	<b>36.78</b>	<b>0.580</b>	<b>0.017</b>	<b>0.666</b>

Despite the range of samples analysed, the calibration of a\* proved to be of less value than L\* with a lower correlation of determination (Table 6.12 R<sup>2</sup>=0.72, SEC=0.378), as can be seen in the diagram of the calibration (Figure 6.10). Hence the prediction of the validation set (Table 6.12 R<sup>2</sup>=0.58) was also lower, achieving an uncorrected SEP of 0.017 and a Bias 0.666 (Table 6.12). As more terms were included into the calibration of CIE a\* both the coefficient of determination and SEC improved (Table 6.12 Term=1, R<sup>2</sup>=0.238, SEC=0.603; Term=6 R<sup>2</sup>=0.72, SEC=0.378). The F-value of the CIE a\* calibration displayed a gradual improvement as more terms were added to the calibration, though the final value is small (Table 6.12 Term=1, F-value= 28.75; Term=6, F-value=36.78). [11]. In contrast to the values obtained for the correlation coefficients of calibration and prediction for L\* which were satisfactory, the value for a\* were low and the predictability less than required for a reliable assessment on a new sample set. This was not as critical as the a\* value, the red to green colour coordinate is not the essential coordinate in sultana colour which is b\*, the yellowness parameter [5].

### 6.3.3 CIE b\*

Table 6.13 Laboratory parameters summary of CIE b\*

# Samples (in calibration)	# Samples in validation	Range	Mean	Standard Deviation
134(107)	27	2.06-12.17	7.67	2.16

#### 6.3.3.1 MLR calibration of CIE b\*

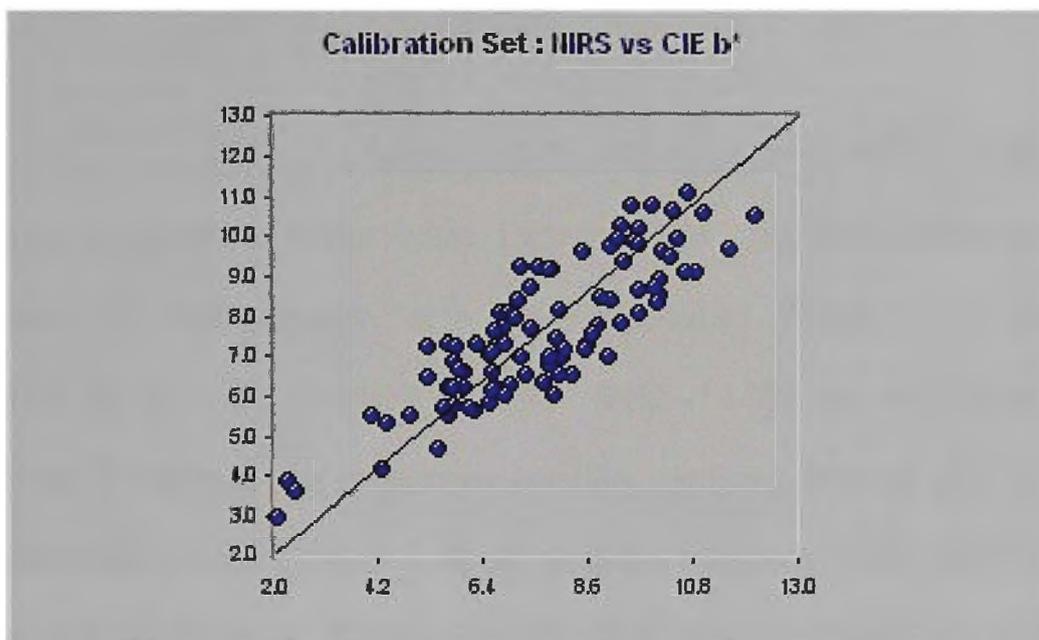


Figure 6.11 A plot of the MLR calibration for CIE b\*

The MLR calibration of CIE b\* showed significant features that originated from the visible region of the spectrum, the most strongly associated frequency (Table 6.14 898nm), however, was attributable to stretching vibrations in methyl groups [9]. The second term (1896nm) was associated with hydroxy stretching vibrations [9] and the fifth (838nm) related to aryl C-H bond stretching vibrations [9,10].

Range and standard deviation of the samples that make up the b\* calibration was larger than previously seen in this chapter, which was indicative of a good calibration. Typically b\* was the easiest of the CIE coordinate calibrations to develop, as its range was relatively large compared to the analytical error of the technique. That is the shade variation in the colours in yellow to light

brown colours was large and it is a critical colour coordinate with respect to the economic value of sultanas.

Table 6.14 Terms contributing to the airstream sorting MLR calibration for CIE b\* (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative).

Terms	Outliers	$\lambda$ (nm)	R <sup>2</sup> (Cal)	SEC	F-value	R <sup>2</sup> (Val)	Bias	SEP
1	11	898	0.485	1.48	85.76	-	-	-
2	"	1898	0.585	1.34	63.49	-	-	-
3	"	498	0.642	1.25	53.27	-	-	-
4	"	622	0.675	1.21	45.77	-	-	-
5	"	838	0.717	1.123	44.10	-	-	-
<b>6</b>	<b>11</b>	<b>530</b>	<b>0.731</b>	<b>1.101</b>	<b>38.96</b>	<b>0.745</b>	<b>0.347</b>	<b>1.98</b>

As more summation terms were included into the calibration both the coefficient of determination and SEC improved (Table 6.14 Term=1, R<sup>2</sup>=0.485, SEC=1.48; Term=6, R<sup>2</sup>=0.731, SEC=1.101). As more terms were added the F-value of the calibration steadily declined, not surprising as the calibration was small and from a single seasons samples (Table 6.14 Term=1, F-value=85.76; Term=6, F-value=38.96) [11]. The validation coefficient of determination achieved was 0.745 with an uncorrected SEP of 1.98 and a Bias of 0.347.

Due to the size of the calibration the correlation achieved (Table 6.14 R<sup>2</sup>=0.731, SEC=1.101) was not as robust as a\* of other research workers, on more uniform samples, such as Flinn *et al* [13]. This fact is illustrated in the provided calibration plot (Figure 6.11). Consequently the calibrations ability to predict the validation set is also not as high (Table 6.14 R<sup>2</sup>=0.745).

Due to the importance of the CIE coordinate b\* for fruit maturity, a calibration using PLS was developed [5].

### 6.3.3.2

### PLS calibration of CIE b\*

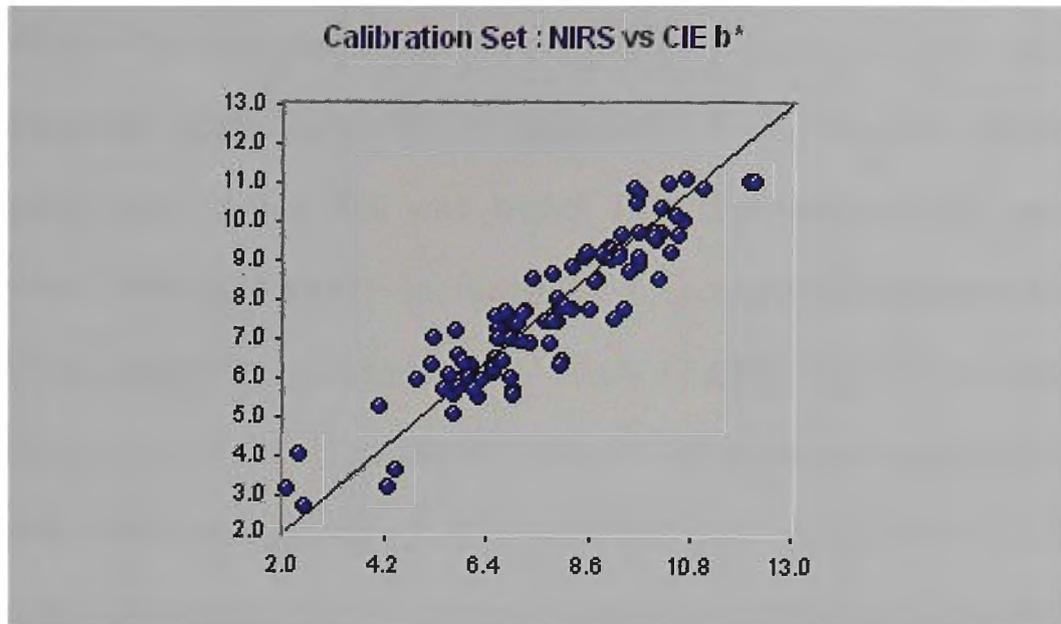


Figure 6.12 A plot of the PLS calibration for CIE b\*

A plot of the calibration set of this calibration is given above (Figure 6.12). PLS more successfully developed a calibration for predicting CIE b\* than MLR as can be seen by comparing the coefficients of determination of both calibrations (Table 6.14 MLR  $R^2=0.731$ , Table 6.15 PLS  $R^2=0.825$ ). As more factors were included both the coefficient of determination and SEC improved rapidly (Table 6.15 Factor=1,  $R^2=0.093$ , SEC=1.98; Factor=10,  $R^2=0.858$ , SEC=0.825).

Table 6.15 Factors contributing to the airstream sorting PLS calibration of CIE b\* (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

PLS								
Factors	Outliers	$R^2$ (Cal)	SEC	SECV	F-value	$R^2$ (Val)	Bias	SEP
1	11	0.093	1.98	2.01	9.30	-	-	-
2	"	0.461	1.53	1.72	38.5	-	-	-
3	"	0.514	1.46	1.59	31.4	-	-	-
4	"	0.548	1.42	1.57	26.7	-	-	-
5	"	0.677	1.21	1.50	36.4	-	-	-
6	"	0.740	1.09	1.43	40.8	-	-	-
7	"	0.783	1.00	1.46	43.7	-	-	-
8	"	0.817	0.92	1.45	47.0	-	-	-
9	"	0.837	0.88	1.47	47.3	-	-	-
<b>10</b>	<b>11</b>	<b>0.858</b>	<b>0.825</b>	<b>1.460</b>	<b>49.53</b>	<b>0.770</b>	<b>0.293</b>	<b>1.68</b>

The trend seen in the F-value of the PLS calibration followed the general improvement that has been seen previously when more factors were added to the calibration (Table 6.15 F=1, F-value=9.30; F=10, F-value=49.53) and reached a final F-value that was higher than that obtained for the MLR calibration (Table 6.14 MLR F-value=38.96), which gave an indication that the calibration progressed well since [11]. SECV steadily improved (Table 6.15 F=1, SECV=2.01; F=10, SECV=1.46) as more factors were introduced to the calibration, thus the predictive model remained reliable even after increasing the number of factors [12]. The resultant validation coefficient of determination was 0.770 with an uncorrected SEP of 1.68 and Bias of 0.293 (Table 6.15).

### 6.3.4 Water activity

Table 6.16 Laboratory parameters summary of Aw

# Samples (in calibration)	# Samples in validation	Range	Mean	Standard Deviation
134(107)	27	0.3915-0.532	0.460	0.0292

#### 6.3.4.1 MLR calibration of Aw

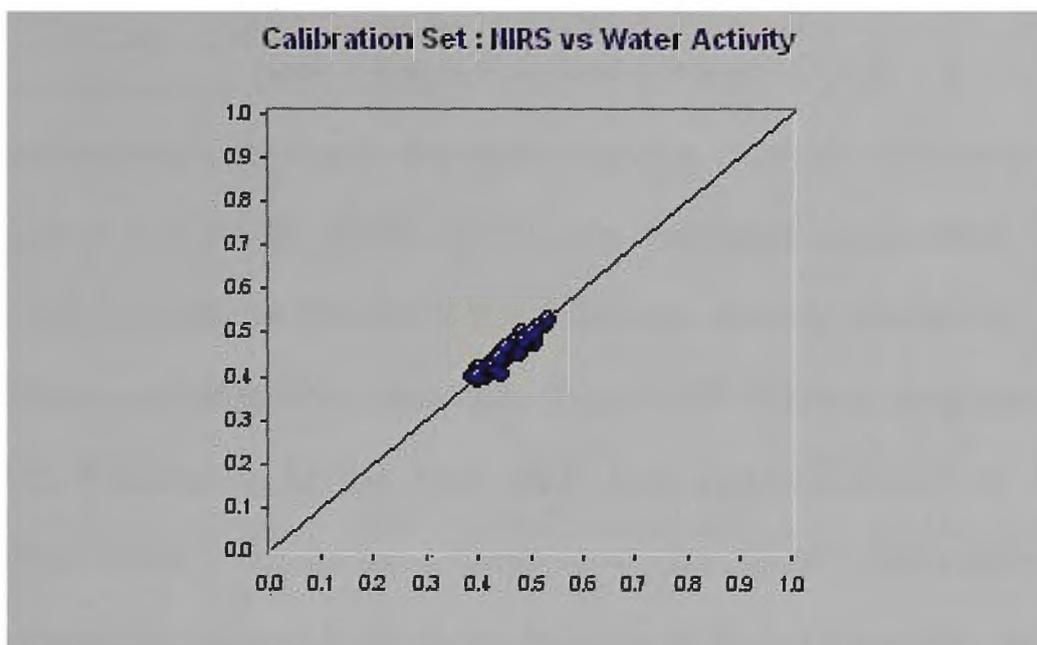


Figure 6.13 A plot of the MLR calibration of water activity

The two most significant terms (1280nm, 1256nm) was attributed to C-H stretching vibrations [9], three terms were attributed to Hydroxyl groups, terms three (1408nm) and six (1840nm) were organic hydroxy stretching vibrations, while term four (2378nm) was an organic hydroxy deformation [9]. Term five (528nm) was within the visible region [8]. The calibration for Water activity was developed by measuring 134 unprocessed samples using a Decagon water activity meter. Typically water activity is the simplest to achieve a reliable calibration, with little or no calibration optimisation required.

Again this water activity calibration proved to be no exception, no outliers were removed from this calibration and a good correlation was achieved (Table 6.17  $R^2=0.923$ ,  $SEC=0.0084$ ) as shown by the calibration plot (Figure 6.13).

Table 6.17 Wavelengths contributing to the air steam sorting MLR calibration of water activity (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

MLR								
No. Terms	Outliers	$\lambda$ (nm)	$R^2$ (Cal)	SEC	F-value	$R^2$ (Val)	Bias	SEP
1	-	1280	0.879	0.010	740.2	-	-	-
2	-	1256	0.905	0.0092	476.5	-	-	-
3	-	1408	0.910	0.0089	338.6	-	-	-
4	-	2378	0.916	0.0087	269.2	-	-	-
5	-	528	0.919	0.0086	223.1	-	-	-
6	-	<b>1840</b>	<b>0.923</b>	<b>0.0084</b>	<b>193.6</b>	<b>0.956</b>	<b>0.0</b>	<b>0.0069</b>

Its reliability has been clearly shown by its ability to predict the validation set (Table 6.17  $R^2= 0.956$ ,  $SEP=0.0069$ ). The validation set showed no Bias (Table 6.17). While the F-value of this calibration sharply decreased as more terms were added to the calibration (Table 6.17 Term=1, F-value= 740.2; Term=6, F-value=193.6) the final value was large, indicative of a valid predictive model [11]. As more terms were introduced to the calibration to compensate for variance between the spectroscopic and laboratory data, both the SEC and coefficient of determination improved (Table 6.17 Term=1,  $R^2=0.879$ ,  $SEC= 0.01$ ; Term=6,  $R^2=0.923$ ,  $SEC=0.0084$ ) also indicating a good predictive calibration. A limited number of studies have used dew point water activity in conjunction with NIRS, despite the near universal application of water activity in food and agricultural fields. Work by CC Huxoll [14,15] developed calibrations for dew point and vacuum oven moisture content of Californian raisins using diffuse reflectance NIRS of whole samples, and showed good potential for the measurement of Aw using NIRS, as a correlation coefficient of 0.979. This is comparable in performance to the calibration developed in this study (Table 6.17 MLR  $R^2=0.923$ ), taking into account the reduced sample population and sample range.

### 6.3.5 Titratable Acidity

Table 6.18 Laboratory parameters summary of Fruit Maturity TA

# Samples (in calibration)	# Samples in validation	Range	Mean	Standard Deviation
134(107)	27	152.95-466.8	226.6	44.84

#### 6.3.5.1 MLR calibration of TA

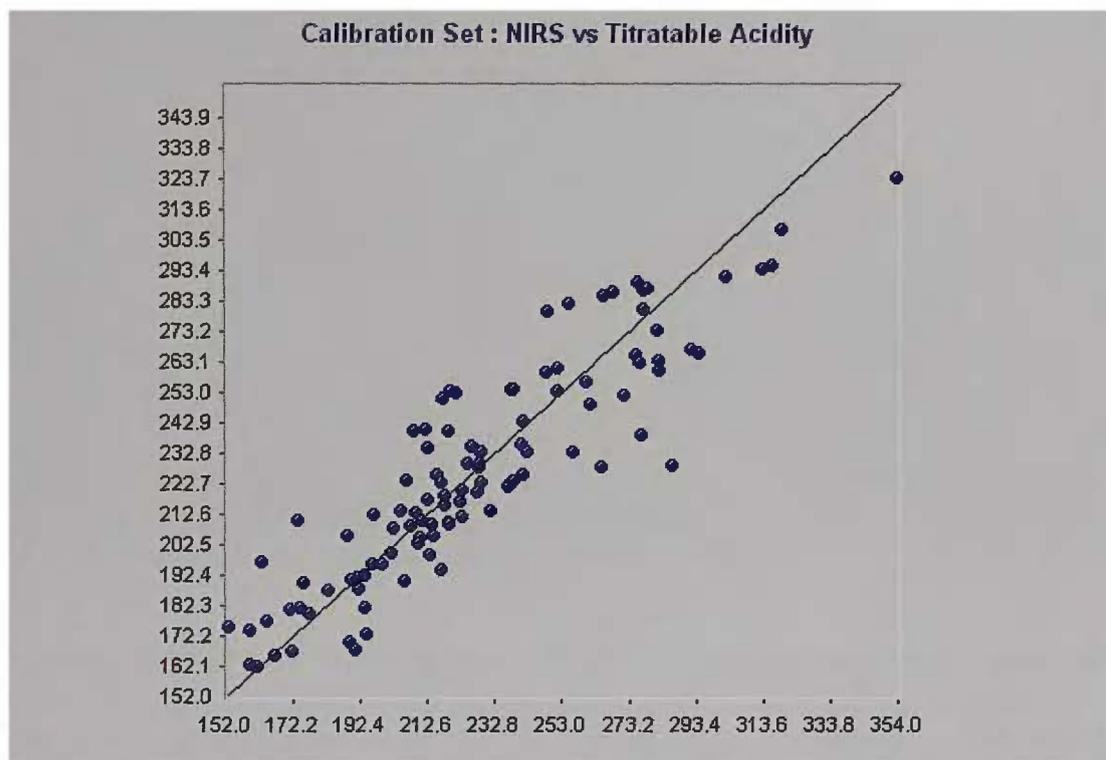


Figure 6.14 A plot of the calibration of Titratable acidity using MLR.

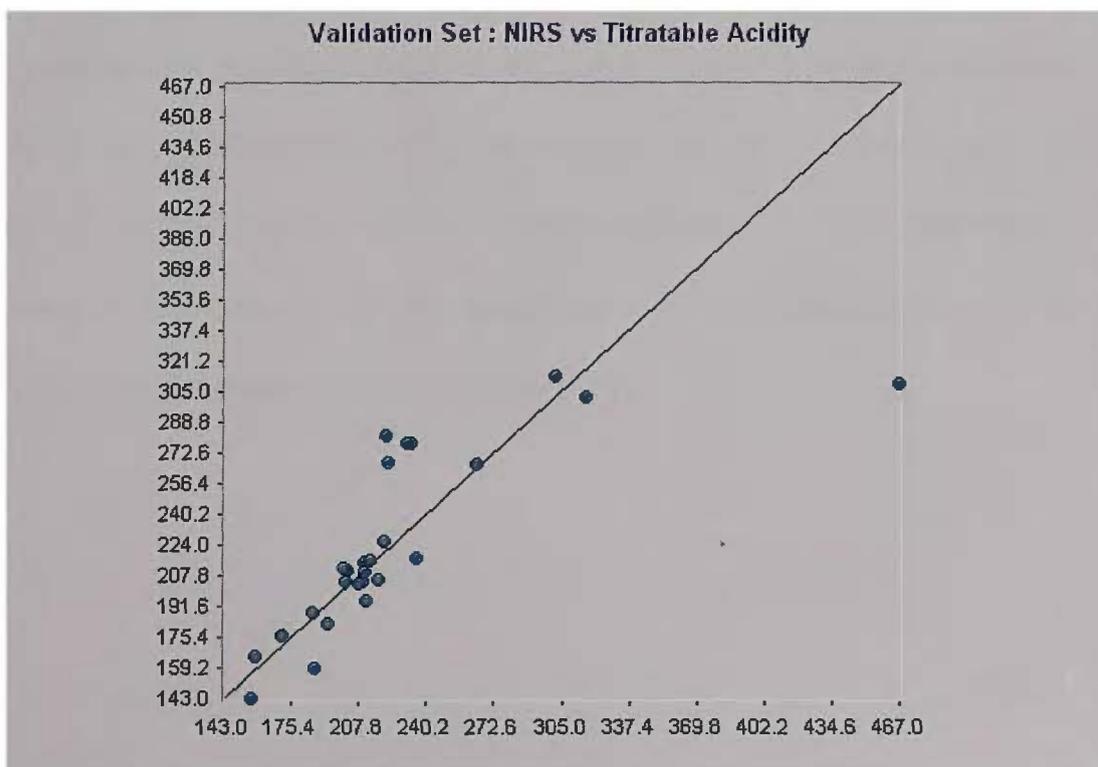


Figure 6.15 A plot of the validation set of titratable acidity using MLR.

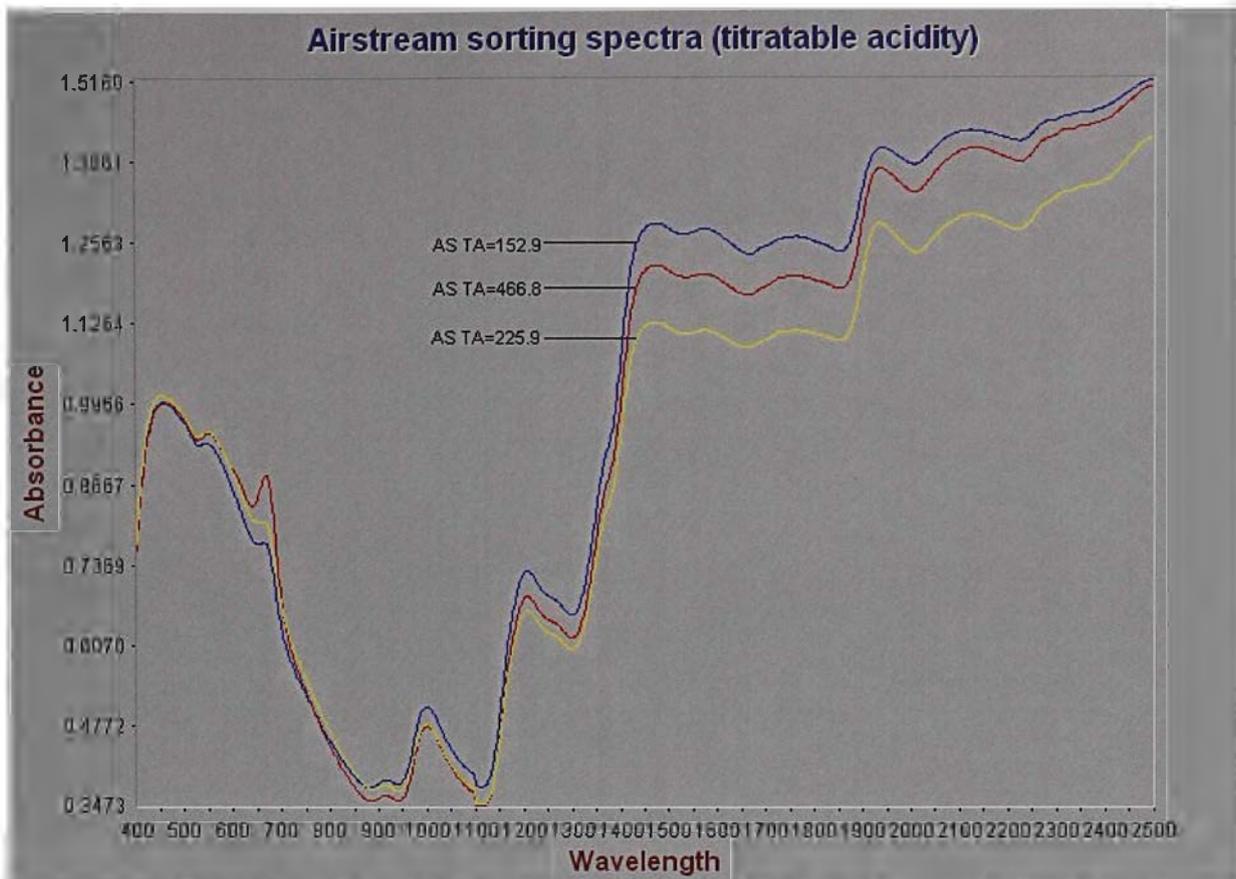


Figure 6.16 Spectra of the minimum (blue), maximum (red) and mean (yellow) samples of titratable acidity

A diagram that shows three spectra from the sample set were seen above (Figure 6.16), these spectra show the whole region covered by the spectrum used in the calibration (400 to 2500 nm) and are the spectrum of the sample that yielded the minimum laboratory value (TA=152.9) the maximum value (TA=466.9) and a spectrum of a laboratory value as close as possible to the mean value (TA=225.9) of the sample set used in this calibration. Also included is a close up of the region of the first wavelength of the MLR calibration of the three spectra (Figure 6.17).

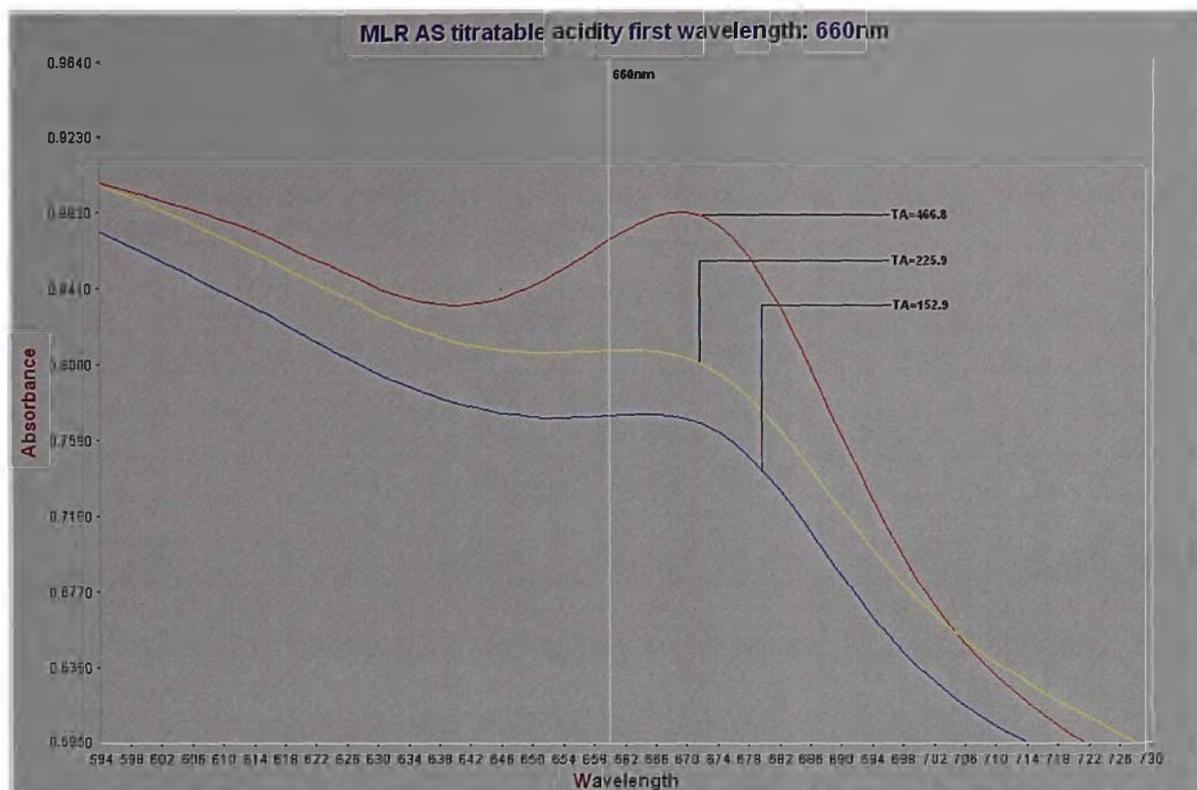


Figure 6.17 Close up of NIR spectra at the first wavelength of the combined MLR calibration

The titratable acidity calibration showed a definite link to the colour of the fruit as the most significant term (Table 6.19 Term 1, 660nm), was located within the visible region [8]. The rest of its terms, however, were not as obvious to assign. Terms two (922nm) and three (1756nm) were located within a region of the spectra that is associated with stretching vibrations of C-H bonds [9]. Term four (2048nm) was linked to N-H organic amine bond stretching vibrations [9,10]. Five (2386nm) was linked to organic hydroxy group stretching vibrations [9] and the last term (1390nm) was attributed to methylene C-H bond stretching and deformation vibrations [9].

The calibration developed during this aspect of the work has correlated well (Table 6.19  $R^2=0.820$ ,  $SEC=17.7$ ) and only two outliers were removed. The calibration developed reliably predicted the samples in the validation set (Table 6.19  $R^2=0.775$ ). During the development of the calibration both the coefficient of determination and SEC improved as more terms were introduced (Table 6.19 Term=1,  $R^2=0.587$ ,  $SEC=26.12$ ; Term=6,  $R^2=0.820$ ,

SEC=17.7). Plots of the MLR calibration set (Figure 6.14) and validation set (Figure 6.15) are shown above.

Table 6.19 Wavelengths contributing to the airstream sorting calibration of titratable acidity (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

MLR								
Terms	Outliers	$\lambda$ (nm)	R <sup>2</sup> (Cal)	SEC	F-value	R <sup>2</sup> (Val)	Bias	SEP
1	2	660	0.587	26.12	145.29	-	-	-
2	"	922	0.705	22.21	120.53	-	-	-
3	"	1756	0.752	20.45	101.13	-	-	-
4	"	2048	0.789	18.96	92.55	-	-	-
5	"	2386	0.810	18.11	83.27	-	-	-
<b>6</b>	<b>2</b>	<b>1390</b>	<b>0.820</b>	<b>17.71</b>	<b>73.43</b>	<b>0.775</b>	<b>-6.40</b>	<b>39.7</b>

The F-value decreased as more terms were added to the calibration, though the final value reached was high which indicated the calibration was robust (Table 6.19 Term=1, F-value=145.29; Term=6, F-value=73.43) [11]. The validation coefficient of determination was 0.775 with an uncorrected SEP of 39.7 and Bias of -6.40 (Table 6.19).

### 6.3.5.2 PLS calibration of TA

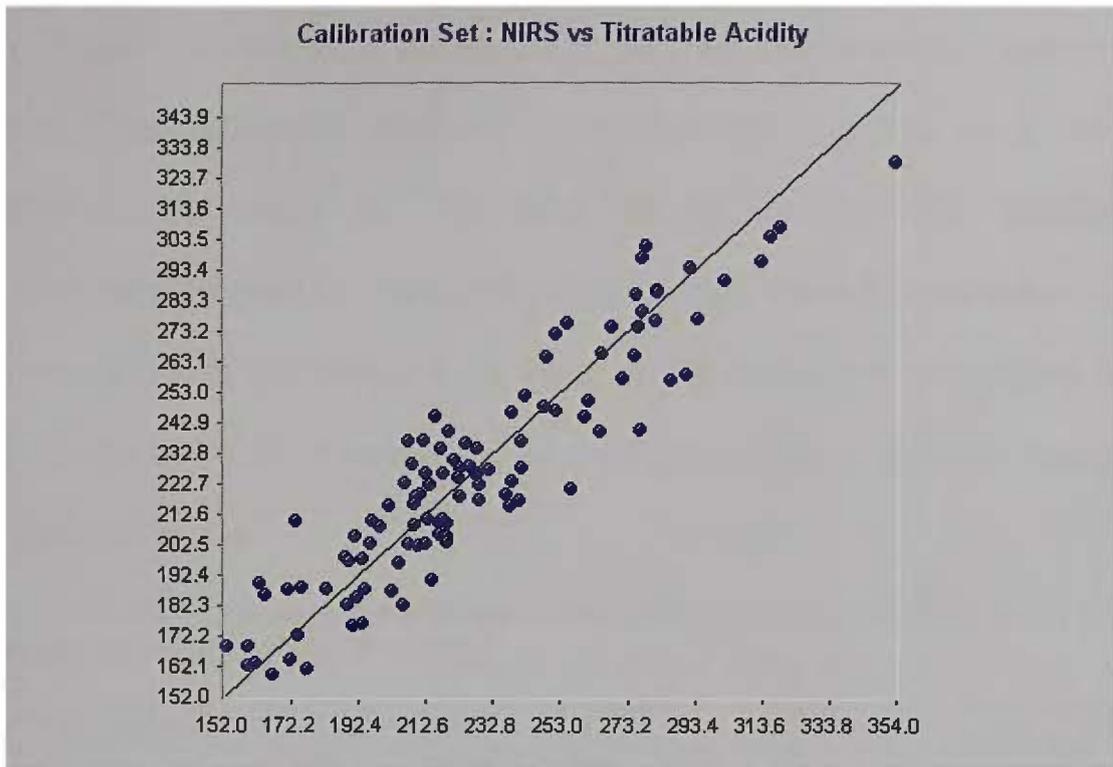


Figure 6.18 A plot of the calibration of titratable acidity using PLS.

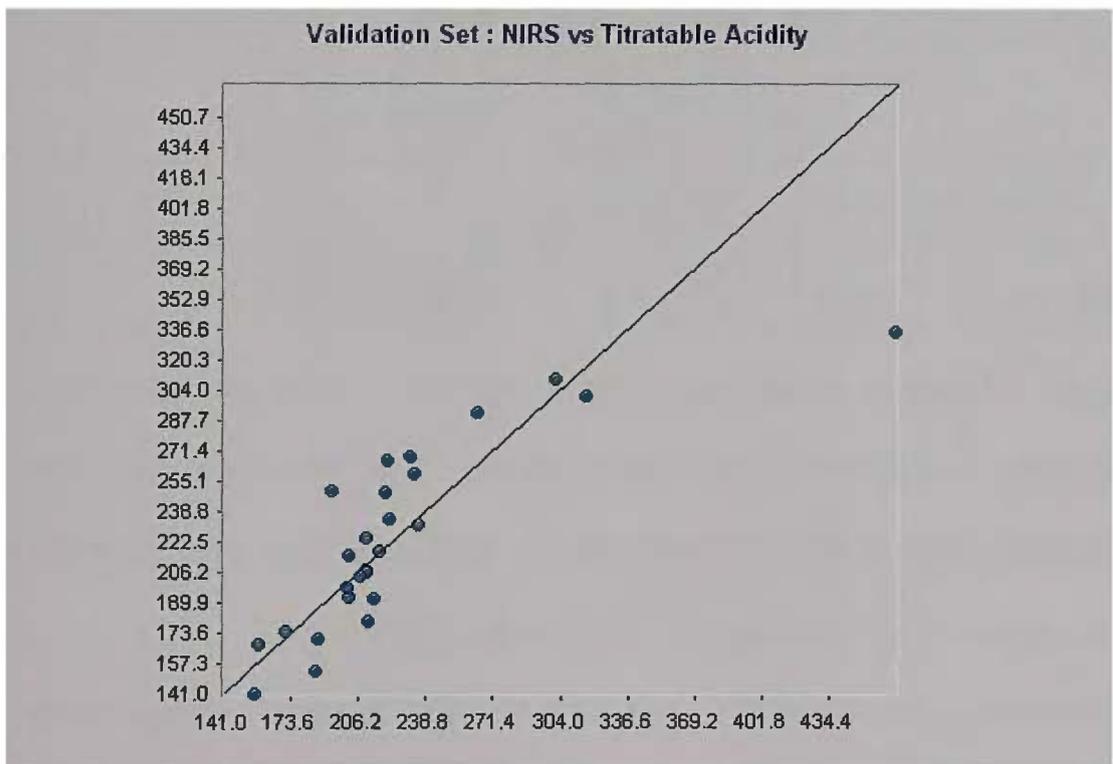


Figure 6.19 A plot of the validation set of titratable acidity using PLS.

The range and standard deviation for the samples used in this calibration were large. For this PLS calibration it was not necessary to remove any outliers. The calibration required 10 factors and resulted in a strongly correlating calibration (Table 6.20  $R^2=0.862$ ,  $SEC=15.83$ ). The reliability of this calibration is good and the ability to predict the validation set proved most successful (Table 6.20  $R^2=0.822$ ). Plots of the calibration set (Figure 6.17) and the validation set (Figure 6.18) of the PLS calibration for titratable acidity are included above.

Table 6.20 Factors making up the airstream sorting PLS calibration of titratable acidity (Pre-treatment: N-point smooth, 2<sup>nd</sup> derivative)

PLS								
Factors	Outliers	$R^2$ (Cal)	SEC	SECV	F-value	$R^2$ (Val)	Bias	SEP
1	-	0.507	28.65	29.06	106.8	-	-	-
2	-	0.584	26.42	27.96	72.45	-	-	-
3	-	0.637	24.82	27.15	59.62	-	-	-
4	-	0.712	22.23	25.33	62.32	-	-	-
5	-	0.750	20.79	24.78	60.09	-	-	-
6	-	0.785	19.38	24.89	60.28	-	-	-
7	-	0.812	18.21	25.06	60.55	-	-	-
8	-	0.825	17.65	25.28	57.32	-	-	-
9	-	0.840	16.99	25.61	55.91	-	-	-
<b>10</b>	-	<b>0.862</b>	<b>15.83</b>	<b>27.45</b>	<b>59.58</b>	<b>0.822</b>	<b>-3.84</b>	<b>35.0</b>

When more factors were introduced to the calibration to explain the variance between the laboratory and spectroscopic data, both the coefficient of determination and SEC improved (Table 6.20  $F=1$ ,  $R^2=0.507$ ,  $SEC=28.65$ ;  $F=10$ ,  $R^2=0.862$ ,  $SEC=15.83$ ), while the SECV and the F-value of the calibration slightly decreased (Table 6.20  $F=1$ ,  $SECV=29.06$ ,  $F\text{-value}=106.8$ ;  $F=10$ ,  $SECV=27.45$ ,  $F\text{-value}=59.58$ ), which indicated that the calibrations cross validation error was improving [12] and the size of the final F-value showed a calibration capable of reliable prediction [11]. The resultant validation coefficient of determination was 0.822 with an uncorrected SEP of 35.0 and a Bias of -3.84.

### 6.3.5.3 Comparison to literature examples

The calibrations developed by McGlone *et al* [16,17] are analogous to the experiments carried out as a part of this study. The techniques used by all researchers in this field are non-destructive: in this particular example there was not any specialised sample presentation as whole fruit is used. In one case Royal Gala apples [16] were used and in the other Satsuma Mandarins [17]. The results obtained by these research projects were typical of whole fruit and lower than that obtained with uniform samples. More successful measuring was made with soluble solids content than titratable acidity, giving a coefficient of determination of 0.65. The reason that more success was obtained in predicting sultana titratable acidity was because the samples that McGlone *et al* used were whole single fruit while this project measured a sample containing several hundred fruit, which would mean the sample is more homogeneous and representative than the single fruit experiments. Navratil *et al* [18] had greater success developing titratable acidity calibrations for yoghurts and filmjölks (yogurt samples  $R^2=0.999$ , filmjolk samples  $R^2=0.989$ ) than the whole fruit work by McGlone *et al* (TA mandarins  $R^2=0.60$ ) [16,17], as the samples they measured were highly homogeneous. Both studies by their nature and the results they obtained contrast those obtained with unprocessed sultanas, as whole unprocessed sultanas are more uniform than the single whole fruit of McGlones studies and the much less uniform highly homogeneous Filmjolk work by Navratil *et al*, and this is reflected in the coefficient of determination obtained (Table 6.19 MLR  $R^2=0.820$ ; Table 7.20 PLS  $R^2=0.862$ ).

## **6.4 Conclusion**

### **6.4.1 Harvest time versus fruit maturity**

The results obtained by airstream sorting indicate that fruit development on a single plant did not occur uniformly. In previous research it has been observed that within even a bunch fruit within a bunch would have developed at different rates dependant on a number of factors, shading and position of fruit in the bunch individual fruit size, non-uniform post-harvest drying or uneven fruit development due to shading contribute to this [9,19,20]. Such variation would result in irregular fruit development ripening, even simply different airstream flow behaviour that would result in a result similar to that observed [1]. The scope of the experiment would have to be increased to verify if the trend was repeatable.

### **6.4.2 Proportion of Fruit Versus Harvest time**

There was a general trend that indicated there was an increase of the most mature fraction over the course of the experiment, the reverse being true for the two most immature fractions. Again it is what occurred between the extremes that is interesting, with relatively high maturity fractions there was observed a general decline in the amount of fruit recovered, while in less mature fractions, an irregular trend of initial increase followed by decline, almost as if fruit are maturing between experimental observations at a far greater rate than other fruit. Decreasing the time scale between the experimental observations and increasing the number of vines examined within the experiment would confirm such observations.

### **6.4.3 Colour versus Harvest time**

Across all the fractions collected an almost uniform trend (refer to Section 6.2.3, Tables 6.3, 6.4 and 6.5) was observed, whether this is linked to variations in arginine levels is beyond the scope of this experiment, but it is likely to be the case [21]. It also can be concluded that a considered choice must be taken by growers to maximise the likelihood of achieving full bodied, light coloured fruit. CIE  $a^*$  followed a very similar trend to  $L^*$ , and the importance in regard to crown grading was by itself low but if used in combination with  $b^*$  can be used as an indicator of fruit type. CIE  $b^*$  however showed quite different behaviour to  $L^*$  or  $a^*$ , and its influence on the grade and type of dried fruit is critical as yellowness is the essential colour parameter for sultanas.

Addressing the experimental hypothesis:

- This series of experiments successfully shown the feasibility of a series of calibrations that would prove applicable to the determination of unprocessed berry maturity.

## 6.5 Bibliography

1. Kasimatis, AN, Vilas, EP Jr., Swanson, FH, Baranek, PB. "Relationship of Soluble Solids and Berry Weight to Airstream Grades", *American Journal of Enology and Viticulture* 1977; 28(1): 8-15.
2. Christensen, LP, Bianchi, ML. "Comparison of Thompson Seedless Clones for Raisin Production", *American Journal of Enology and Viticulture* 1994; 45(2): 150–154.
3. Christansen, LP, Bianchi, ML, Lynn, CD, Kasimatis, AN, Miller, MW. "The Effects of Harvest Date on Thompson seedless Grapes and Raisins (I) Fruit Composition, Characteristics and Yield", *American Journal of Enology and Viticulture* 1995; 46(1):10-16.
4. Christansen, LP, Bianchi, ML, Miller, MW, Kasimatis, AN, Lynn, CD. "The Effects of Harvest Date on Thompson seedless Grapes and Raisins (II) Relationships of Fruit Quality Factors", *American Journal of Enology and Viticulture* 1995; 46(4):493-498.
5. Uhlig, BA, Clingeleffer, PR. "Influence of Grape (*Vitis Vinifera* L.) Berry Maturity on Dried Fruit Colour" *Journal of Horticultural Science and Biotechnology* 1998; 73(3): 329-339.
6. Mark, H. *Principles and practice of spectroscopic calibration*. Brisbane: J.Wiley and sons, Inc; 1991;p.104
7. *Near Infrared technology in the agricultural and food industries*. Second Edition. Edited by P Williams and C Norris. St. Paul (Mi): American Association of Cereal Chemists; 2001;p.153-154

8. Osborne, B, Fearn, T, Hindle, P. *Near Infrared Spectroscopy with applications in food and beverage analysis*. 2<sup>nd</sup> Edition. Singapore: Longman Scientific and Technical; 1993.
9. The Chart '*Near Infrared Adsorptions: Major analytical bands and their relative peak positions*'. Silver Spring (MD): NIRSystems, Inc; 1992.
10. *Near Infrared technology in the agricultural and food industries*. Second Edition. Edited by P Williams and C Norris. St. Paul (Mi): American Association of Cereal Chemists; 2001;p.33-36.
11. *Near Infrared technology in the agricultural and food industries*. Second Edition. Edited by P Williams and C Norris. St. Paul (Mi): American Association of Cereal Chemists; 2001p.151-152, 180 194, 196, 212.
12. *Near Infrared technology in the agricultural and food industries*. Second Edition. Edited by P Williams and C Norris. St. Paul (Mi): American Association of Cereal Chemists; 2001; p.154
13. Flinn, PC, Black, RG, Iyer, L, Brouwer, JB, Meares, C. "Estimating the Food Processing Characteristics of Pulses by Near Infrared Spectroscopy, using Ground or Whole Samples", *Journal of Near Infrared Spectroscopy* 1998;6:213-220.
14. Huxoll, CC. "Near infrared analysis potential for measuring the water activity of raisins". *Conference Proceedings of the IFT Annual Meeting*. 1995; p.70.
15. Huxoll, CC. "Assessment of Near Infrared (NIR) Diffuse Reflectance analysis for measuring Moisture and Water Activity in Raisins", *Journal of Food Processing and Preservation* 2000;24(4): 315-333.

16. McGlone, VA, Fraser, DG, Jordan, JB, Kunnemeyer, R. "Internal Quality Assessment of Mandarin Fruit by Visible-Near Infrared Spectroscopy", *Journal of Near Infrared Spectroscopy* 2003;11(5):323-332.
17. McGlone, VA, Jordan, RB, Martinsen, PJ. "Visible-NIR Estimation at Harvest of Pre- and Post-storage Quality Indices for Royal Gala Apple", *Post-harvest Biology and Technology* 2002;25(2):135-144.
18. Navratil, M, Cimander, C, Mandenius, CF. "Online Multi-Sensor Monitoring of Yoghurt and Filmjolk Ferments on Production Scale", *Journal of Agricultural and Food Chemistry* 2004;52(3):415-420.
19. Uhlig, BA. "Effects of solar radiation on grape (*Vitis Vinifera* L.) composition and dried fruit colour", *Journal of Horticultural Science and Biotechnology* 1998 73(1) 111-123.
20. Uhlig, BA, Clingeleffer, PR. "Ripening Characteristics of the Fruit From *Vitis Vinifera* L. Drying Cultivars Sultana and Merbein Seedless under Furrow Irrigation" *American Journal of Enology and Viticulture* 1998 49(4) 375-382.
21. Frank, D. *Investigation of the biochemical basis of browning during the storage of sultanas, a PhD Thesis*. Victoria University, Werribee Campus; 2001.

## **7.0 General Discussion and Overall Conclusions**

### **7.1 Assessment of processed sultana quality using NIRS**

The results presented in Chapters 4 and 5, the two results chapters, Part 1 and Part 2, respectively, clearly demonstrate an ability to assess a variety of quality factors of processed sultanas. In all, except one parameter, calibrations were successfully developed and showed sufficient reliability to be used in an industrial setting. The combined seasons calibrations for CIE L\*, the lightness coordinate, for both MLR and PLS will now be given. Firstly, the parameters for MLR statistics: Coefficient of determination,  $R^2$  is 0.840; Standard Error of Calibration, SEC, is 0.780 and the Validation  $R^2$ , on unknown samples is 0.840. Secondly, the PLS statistics are:  $R^2 = 0.835$ , SEC = 0.727 and the Validation,  $R^2 = 0.792$ . These determination coefficients are in an acceptable range for an agricultural product. Considering the importance of CIE L\* in relation to dried fruit Crown Grades as discussed by Grncarevic and Lewis [1], in their article and also the work by Uhlig and Clingeffer [2], the development of a successful calibration for the L\* parameter is of considerable value to the industry.

CIE a\*, the redness to greenness colour coordinate, combined seasons calibrations for MLR are as follows:  $R^2=0.777$ , SEC=0.262 and validation  $R^2=0.794$ . For the PLS calibration, the same parameters are:  $R^2=0.787$ , SEC=0.275 and validation,  $R^2=0.789$ , also showed the ability to predict successfully. It should be noted that the a\* correlation value is not so critical with respect to an essentially yellow to light brown sultana. Hence, the correlation coefficients are not as high as other calibrations developed in this

work. The  $a^*$  coordinate is important in identifying fruit that is of lesser quality as a red or green tinge to sultanas is an undesirable colour and reduces the economic value of the fruit.

CIE  $b^*$  (the yellowness coordinate) calibrations gave excellent predictions of fruit quality. For MLR the statistics are as follows:  $R^2=0.869$ ,  $SEC=0.666$  and validation  $R^2=0.932$ . Considering PLS the equivalent values are:  $R^2=0.884$ ,  $SEC=0.630$  and the validation,  $R^2=0.924$ . This parameter is an important one for predicting sultana quality as it, gives a measure of the gold colour of the sultanas which is the colour that brings the highest price on the international and local markets. It is of considerable benefit to the industry to identify quickly the premium grade sultanas. It should also be mentioned that these three colour coordinates are simultaneously measured when the sultanas are assessed and the values of the  $L^*$ ,  $a^*$  and  $b^*$  in combination are used to distinguish different grades of sultanas.

The combined water activity ( $A_w$ ) calibration gave the highest correlation coefficients of all those developed. As can be seen below both statistical development methods displayed high correlations for both the calibration and validation data sets. Considering the MLR statistics first:  $R^2=0.936$ ,  $SEC=0.0059$  and validation,  $R^2=0.965$ . There was little difference in the numerical values for these parameters with PLS:  $R^2=0.949$ ,  $SEC=0.0053$ , validation  $R^2=0.957$ . This indicates that both methods for calibration method development are of equal standing. The importance of this constituent is not as important for processed fruit compared to unprocessed fruit, as any Maillard browning that has occurred would have already affected the fruit. The value of the water activity calibration lies in using the NIRS as a means of

feedback for the processing line, as the fruit are washed during the process, then dried. The ability to check water levels within fruit allows the processor to make informed decisions about time spent in drying the fruit in and/or the temperature which allows an optimum fruit mass, free of excess moisture weight, which would reduce transport expenses.

Titrateable Acidity as a constituent for NIRS calibrations, despite the delay between harvest and processing affecting the levels of fruit acids present, showed a clear correlation to the NIR spectra. It was thought that this constituent may prove difficult to develop since a calibration as the time between harvesting and processing could result in the fruit ripening further in storage, causing low levels of fruit acids with little content variation: a classically difficult set of circumstance to develop a chemometric calibration. This may have been the reason that the lower MLR calibration values:  $R^2=0.736$  SEC=14.5 and the validation  $R^2=0.70$ , was lower than ideal for use as a linear regression technique. PLS, however, demonstrated a far more robust performance, in developing a predictive model for TA. The statistical values being:  $R^2=0.877$ , SEC=10.33 and validation  $R^2=0.798$ . The importance of titrateable acidity for processed sultanas is less than that for unprocessed sultanas, as the fruit may have lost some of the fruit acids in storage, however it is an indication of the taste of the fruit, and therefore may be used in future as a tangible measure of fruit eating quality. Titrateable acidity of fruit would also give some indication of fruit maturity at picking and a reinforcing indication of the presence or absence of green tinge fruit within the batch.

The combined seasons calibration for Kjeldahl nitrogen, total nitrogen content, demonstrated an ability to reliably predict this constituent in processed

sultanas. This constituent is a measure of the amino acids in sultanas, and the uptake of nitrogen from the soil and fertilizer employed; arginine and proline being the major amino acids present. The standard deviation between samples would have a correlation with the amount of fertilizer used and soil type [3,4,5] and should be diverse. Strongly correlating calibrations were developed for both MLR and PLS giving reliable predictive models. The MLR statistics are:  $R^2=0.793$ ,  $SEC=0.180$  and validation  $R^2=0.759$ . For PLS:  $R^2=0.859$ ,  $SEC=0.147$  and validation  $R^2=0.764$ . The importance of this constituent, while not as crucial as in unprocessed sultanas due to its links to Maillard browning, it is still an important parameter for nutritional value, and will give an indication of future discolouration of processed fruit during storage. [3]

Of all the parameters, Total Lipids proved elusive with respect to developing a successful calibration. This should not be unexpected, as the constituent should conceptually vary little between samples, due to the fact that the dressing oil is evenly dispersed throughout the sample and the natural surface lipids should vary only because of variations in fruit surface area. The low correlation coefficients are therefore understandable, MLR,  $R^2=0.68$  and PLS  $R^2=0.586$ . Due to the poor correlations obtained in the first season, this part of the project was abandoned.

## **7.2 Maturity assessment of unprocessed sultanas using NIRS**

In a departure from the main thrust of the thesis, as a way of answering requests from the industry to devise a new way of assessing fruit by a parameter more meaningful to the true eating quality of sultanas, an assessment of berry maturity was undertaken. As previously stated in this thesis berry maturity of “natural” sultanas is determined by the use of an airstream sorter which separates the fruit by means of a strong airflow into different fractions of different maturities of fruit by varying the airflow rate. As it proved to be conceptually difficult to develop a way of directly relating airstream sorting values into a linear regression model, it was decided to use this technique as a means of artificially creating varying levels of fruit maturity, In addition, fruit from trial plots was picked at different intervals during fruit development to vary fruit maturity, which would result in a set of samples with a broad range of titratable acidity values. This strong relationship between titratable acidity, maturity and Airstream sorting values has already been established by a number of research projects [2,6,7,8,9]. With an established link between titratable acidity and fruit maturity in mind, a calibration developed to predict titratable acidity could be used to predict relative maturity and “full-bodied-ness” of unprocessed fruit. Using the combined samples of the harvest time variation study combined with airstream sorted samples and a series of typical, unsorted samples, a titratable acidity calibration was successfully developed, both MLR and PLS (MLR  $R^2=0.820$ ,  $SEC=0.0084$  validation  $R^2=0.775$  PLS:  $R^2=0.862$   $SEC=15.83$  validation  $R^2=0.822$ ) proved capable of reliably predicting titratable acidity of processed samples. The importance of this calibrations success is clear, it would now be possible to

gauge the maturity and “full-bodied-ness” of fruit that are received by the packing sheds, which should reward growers for waiting sufficient time before harvesting fruit.

### **7.3 Theoretical implications**

The underlying thrust of this project was to apply existing methodologies to a series of industrial problems with the aim of overcoming them. The theoretical outcomes from this project were that a series of parameters are rapidly assessed by a single analytical technique of a non-homogenous sample. Developing this technique to assess sultanas, both processed and unprocessed has required the development of unique mathematical models, which link the analytes or physicochemical parameters to trends seen in spectroscopic data. The use of NIRS to assess quality of processed sultanas is a new application of existing methodologies and technologies, therefore this project is a completely new application NIRS to a subject that has received little attention by this technique.

## 7.4 Practical implications

The follow on effects of this research of successfully developing a NIRS assessment of Processed fruit:

- Now the industry can price the products they develop based on tangible quality parameters.
- The parameters CIE L\*, a\* and b\*, water activity, Kjeldahl protein, and titratable acidity can be reproducibly measured and replicated.
- The requirements of buyers will be more easily met, because the NIRS grading techniques can discriminate between samples more readily than a visual grading based on colour uniformity.
- The grading system reflects more precisely the maturity and taste of the fruit rather than a system based purely on colour.

Maturity assessment of unprocessed fruit by NIRS:

- The growers will be awarded for delivering mature, full-bodied fruit, which will offset the risks involved with supplying such fruit, rather than be awarded for supplying light coloured fruit regardless of maturity.
- The reduction of green tinge, low maturity fruit present in the final product will improve overall fruit quality. This will be achieved by consistently providing a light, more uniformly coloured product that will attract a premium price.
- A reputation for quality fruit will allow Sunraysia sultanas to have a competitive edge over their competitors.

A representative of Foss the Instrument Company examined the packing facilities at Sunbeam Foods. Due to the cost of introducing a dedicated “on-line” instrument at a number of sites in the Mildura district, the seasonal

nature of its utility, and the relatively low throughput volumes, an “at-line” approach was considered more feasible by the industry.

## **7.5 Future research**

With the aims to further pre-empt regulatory requirements to label food products, future applications of NIRS technology would be focused towards developing commonly used parameters of food products, indicating the nutritional impact of the product on consumers. Examples could be common nutritional parameters such as digestible or indigestible fibre, crude ash content and specific mineral levels, sugar levels, and fat content as both a means of informing consumers about the use of vegetable oils to improve the fruits flow characteristics and naturally occurring surface lipids, and as a means of feedback to the processing line to ensure the correct levels of oil are being added. Furthermore, newer spectroscopic techniques, such as FT-NIR or diode array detectors, could also be used to improve the characteristics of the NIR spectrophotometer adopted by the industry. FT-NIR instruments are advantageous over conventional scanning NIR Spectrometers, as they contain fewer moving parts and are therefore conceptually more reliable. Alternatively, simplified instruments that are more portable and cheaper, diode array, with the potential of the instrument being used for limited applications. For example, a simplified instrument capable of measuring water activity levels in drying fruit. This would enable growers to choose the optimum time to transport the fruit to the processor, or manage rain damage of drying fruit more successfully. Such portable instruments could also be used to analyse nitrogen uptake in fruit to identify such problems at their source and therefore

make subtle localised adjustments to growing practices, to limit or completely nullify such problems. Such a use of the technology is conceptually more powerful as it potentially enables the grower to more effectively counter the problems in advance.

## 7.6 Bibliography

1. Grncarevic, M, Lewis, W. "External Colour of Dried Sultanas", *Food Technology in Australia* 1973 November: 562-565.
2. Uhlig, BA, Clingeleffer, PR. "Influence of Grape (*Vitis Vinifera* L.) Berry Maturity on Dried Fruit Colour" *Journal of Horticultural Science and Biotechnology* 1998; 73(3): 329-339.
3. Frank, D. *Investigation of the biochemical basis of browning during the storage of sultanas*, a PhD Thesis. Victoria University, Werribee Campus; 2001.
4. Zerihun, A, Treeby, MT. "Biomass Distribution and Nitrate Assimilation in Response to N Supply for *Vitis Vinifera* L. CV. Cabernet Sauvignon on Fire Vitis Rootstock Genotypes", *Australian Journal of Grape and Wine Research*. 2002;8(3):157-162.
5. Treeby, MT. "Sultana Fruitfulness and Yield: Responses to Rootstock and Nitrogen Supply", *Australian Journal of Experimental Agriculture* 2001;41(5): 681-687.
6. Kasimatis, AN, Vilas, EP Jr., Swanson, FH, Baranek, PB. "Relationship of Soluble Solids and Berry Weight to Airstream Grades", *American Journal of Enology and Viticulture* 1977; 28(1): 8-15.
7. Christensen, LP, Bianchi, ML. "Comparison of Thompson Seedless Clones for Raisin Production", *American Journal of Enology and Viticulture* 1994; 45(2): 150–154.
8. Christansen, LP, Bianchi, ML, Lynn, CD, Kasimatis, AN, Miller, MW. "The Effects of Harvest Date on Thompson Seedless Grapes and Raisins (I) Fruit Composition, Characteristics and Yield", *American Journal of Enology and Viticulture* 1995; 46(1):10-16.

9. Christansen, LP, Bianchi, ML, Miller, MW, Kasimatis, AN, Lynn, CD.  
“The Effects of Harvest Date on Thompson seedless Grapes and Raisins (II) Relationships of Fruit Quality Factors”, *American Journal of Enology and Viticulture* 1995; 46(4):493-498.

## Appendix 1: season one laboratory data

Sample	L	a	b	WA	Nitrogen	Lipids	TA
1a	34.3200	5.9900	8.0400	0.5840		1.4477	
2a	32.9600	5.3500	4.9300	0.5850		1.2361	
3a	35.4700	5.5600	7.5100	0.5675		1.2285	
4a	33.8700	5.4000	8.2700	0.5500		1.6243	
5a	32.1900	5.6200	6.8900	0.5740		1.4058	
6a	33.7000	6.3100	7.9400	0.5530		1.2536	
7a	35.8400	5.5700	7.2900	0.5470		1.4041	
8a	34.0900	5.7800	6.1000	0.5650		1.3075	
9a	34.3300	5.1100	1.0900	0.4960		1.1379	
10a	32.0900	5.8400	6.2900	0.5740		1.4731	
11a	30.9300	6.1100	1.9100	0.5870		1.4058	
12a	34.1300	6.0000	7.3200	0.5675		1.4417	
13a	34.6300	4.4600	9.5000	0.4895		1.5577	
14a	34.6600	4.9400	8.6300	0.5535		1.3943	
15a	33.6400	4.6200	7.9300	0.5630		1.5056	
16a	27.2300	3.0900	3.2900	0.5330		1.2816	
17a	31.6300	4.9000	1.0100	0.5530		1.0436	
18a	33.6100	5.5200	5.7800	0.5860		1.1989	
19a	33.9900	5.1800	7.4800	0.5965		1.3450	
20a	34.0800	4.8400	6.3100	0.5735		1.2212	
21a	37.1700	4.9500	9.1700	0.5240		1.4106	
22a	34.9300	4.9700	8.9000	0.5490		1.1902	
23a	34.6600	4.9300	6.4800	0.5345		1.4444	
24a	34.2600	5.1300	7.9600	0.5670		1.4341	
29-a	34.0400	4.9100	8.0100	0.5510		1.1973	
30-a	34.5400	4.8800	8.3600	0.5420		1.2483	
31-a	33.6400	4.8700	7.9800	0.5690		1.1326	
32-a	34.5800	4.9600	9.0700	0.5310		1.1351	
33-a	33.8800	4.7400	8.9500	0.5560		1.5194	
34-a	34.3500	5.2600	8.4100	0.5555		1.1547	
35-a	32.2700	5.0700	7.3300	0.5580		1.4421	
36-a	34.1600	5.3600	7.6600	0.5345		1.2372	
37-a	34.5400	5.3300	8.4100	0.5350		1.1666	
38-a	33.7700	5.3200	8.4800	0.5540		1.2785	
39-a	35.0600	5.1900	8.6300	0.5375		1.1883	
40-a	33.9000	4.6800	8.6400	0.5605		1.3961	
41-a	34.5900	5.1600	8.0400	0.5420		1.2620	
42-a	36.2000	5.2600	9.2100	0.5425		1.3364	
43-a	35.9200	5.3100	8.9500	0.5255		1.1974	
44-a	34.7900	5.1000	7.0300	0.5000		1.0506	
45-a	35.7000	5.2200	7.5300	0.5075		1.1872	
46-a	36.7700	5.5000	9.7800	0.4845		1.0992	
47-a	31.5100	5.5000	4.1800	0.5465		1.0510	
48-a	30.7300	6.0200	3.9800	0.5610		1.3030	
49-a	31.2900	6.1600	4.5500	0.5510		1.2842	
50-a	32.0200	6.2300	3.9700	0.5550		0.9279	
51-a	31.5600	6.2000	3.8100	0.5610		1.4766	

52-a	31.9900	5.7400	3.6200	0.5410		1.3029
53-a	33.9500	5.1900	5.9900	0.5125		1.4183
54-a	30.8700	6.0800	3.9300	0.5570		1.4058
55-a	35.9800	4.5100	8.3900	0.5045		1.3515
56-a	34.1200	5.1200	7.5900	0.5335		1.5435
57-a	30.8900	5.6000	4.7500	0.5315		1.5797
58-a	35.1900	5.6600	7.1500	0.5010		1.4160
59-a	34.1100	5.7000	6.6200	0.5170		1.4695
60-a	31.4200	6.1700	3.7800	0.5505		1.1623
61-a	33.6000	5.6500	6.4100	0.5440		1.2703
62-a	30.9200	5.7200	4.0500	0.5575		1.2542
63-a	34.2500	5.1000	7.0000	0.5215		1.2635
64-a	32.3300	5.3800	4.0500	0.5555		1.4733
65-a	34.8600	5.5400	7.3200	0.5025		1.3834
66-a	32.8800	5.2900	4.8800	0.5515		1.3057
67-a	32.0600	6.1000	3.9800	0.5785		1.3530
68-a	31.6700	5.4100	4.8700	0.4835		1.3949
69-a	32.4200	5.2900	4.4700	0.5465		1.1593
70-a	33.3700	5.0500	7.0900	0.5125		1.5790
71-a	30.9900	6.0900	3.7900	0.5705		1.3514
72-a	30.4000	5.6700	3.8400	0.5520		1.6056
73-a	30.9100	5.4300	4.4000	0.5350		1.4841
74-a	31.9700	5.4200	4.3600	0.5345		1.3776
75-a	31.6500	5.3700	3.6300	0.5470		1.6542
76-a	31.1500	5.2400	4.3100	0.5675		1.5649
77-a	31.8400	5.5300	4.7100	0.5585		1.3854
78-a	30.2000	5.6300	3.5200	0.5555		1.3638
79-a	31.5200	5.5700	4.5800	0.5345		1.3027
80-a	31.3300	5.2700	4.0000	0.5085		1.6393
81	30.6700	5.1100	5.4100	0.5380	1.8900	217.0000
82	30.4800	5.3000	4.4600	0.5230	2.4500	241.0000
83	32.4700	5.4400	5.3400	0.5280	2.0900	234.0000
84	34.3500	5.9800	5.5600	0.5800	2.1300	221.0000
85	32.8400	5.2500	6.8700	0.5130	2.3200	241.0000
86	34.1700	5.2300	7.1800	0.5650	1.9600	234.0000
87	31.4200	5.4800	4.2200	0.5200	2.2700	206.0000
88	32.5000	5.0900	7.6500	0.5370	1.9000	212.0000
89	34.8600	5.4700	8.6200	0.5770	1.9600	209.0000
90	35.9600	5.4100	8.3100	0.5930	2.0400	215.0000
91	32.7700	5.6300	6.2900	0.5650	1.7400	216.0000
92	35.4200	5.5100	7.8600	0.6040	2.0100	213.0000
93	31.9800	4.4600	8.0900	0.5790	1.9800	234.0000
94	34.2000	5.7300	5.9400	0.5360	1.7100	199.0000
95	33.6000	4.7900	6.3300	0.5800	2.3700	253.0000
96	33.5500	4.8700	7.4600	0.5630	2.0100	235.0000
97	32.4100	5.3900	6.9000	0.5700	1.9900	223.0000
98	33.6700	5.7500	7.6800	0.5320	2.0400	220.0000
99	30.0200	4.7900	5.5100	0.5280	2.5600	281.0000
100	35.1200	5.7400	7.9000	0.5970	2.1100	208.0000
101	33.7800	5.6300	6.7000	0.5310	2.3200	225.0000
102	33.2600	5.5100	4.6300	0.5690	1.9400	232.0000

103	33.1100	5.9600	4.9400	0.5380	3.5000	207.0000
104	30.7200	6.3600	6.8900	0.5650	3.3000	210.0000
105	34.2300	5.7600	5.2900	0.5290	2.9000	254.0000
106	33.8000	5.2400	8.5900	0.5220	3.5000	217.0000
107	30.9700	5.8300	4.5900	0.5525	1.7800	227.0000
108	32.2800	6.1700	5.8300	0.5505	3.4500	206.0000
109	33.3700	5.5300	5.9200	0.5220	2.5900	230.0000
110	31.5800	5.6700	5.6900	0.5440	2.4000	200.0000
111	33.3000	5.6800	7.2600	0.5510	2.4700	216.0000
113	32.6600	5.8200	5.8100	0.5150	1.3100	223.0000
114	29.7800	5.9100	4.7700	0.5735	2.8300	237.0000
115	31.4200	6.2700	5.2600	0.5390	2.9500	240.0000

## Appendix 2: Season two laboratory data

Sample	L	a	b	wa	Nitrogen	TA
1	33.8600	5.0500	6.8100	0.5410	2.0300	213.2000
Cc-b0002	31.2700	4.3100	5.5400	0.5270	3.3000	283.8000
Cc-b0003	31.6000	4.7700	5.3800	0.5700	3.0000	264.8000
Cc-b0004	31.6000	4.9700	8.2500	0.4970	2.4400	211.9000
Cc-b0005	35.4900	4.3300	8.3200	0.5160	2.5300	242.6000
Cc-b0006	36.3900	4.7900	8.3200	0.5330	2.0600	244.8000
Cc-b0007	34.9000	3.8600	9.9200	0.5290	2.1800	241.3000
Cc-b0008	34.7600	3.7300	9.5800	0.5220	2.3800	243.3000
Cc-b0009	35.9500	4.1200	9.2700	0.5320	2.1800	222.4000
Cc-b0010	34.4300	4.2000	8.3000	0.5250	2.6300	241.4000
Cc-b0011	36.0600	4.7400	9.1100	0.5200	2.1300	218.8000
Cc-b0012	34.7700	4.5000	9.5300	0.5260	1.8400	201.5000
Cc-b0013	35.1700	4.5400	8.1100	0.5330	2.5100	203.5000
Cc-b0014	35.5300	5.0400	7.9300	0.5170	2.4700	218.8000
Cc-b0015	36.1300	4.6800	8.3600	0.5400	2.3200	220.0000
Cc-b0016	35.8100	4.8700	8.0300	0.5210	2.4100	246.4000
Cc-b0017	35.6400	5.2800	7.1300	0.5290	1.9500	226.6000
Cc-b0018	35.3900	3.9500	9.6800	0.5250	1.9000	215.2000
Cc-b0019	36.6000	4.9600	9.4100	0.5300	2.0600	238.6000
Cc-b0020	34.6700	4.5200	7.8800	0.5370	2.9300	240.6000
Cc-b0021	35.8000	4.4200	8.1500	0.5410	2.7800	227.3000
Cc-b0022	35.0600	4.5400	8.1700	0.5320	2.2700	255.4000
Cc-b0023	34.2000	4.9800	6.5200	0.5350	2.5300	210.0000
Cc-b0024	34.4300	5.1800	6.9400	0.5400	2.2900	198.7000
25	34.4900	4.4700	8.2000	0.5730	2.2300	191.8000
Cc-b0026	35.9100	4.2500	8.4800	0.5520	2.1900	209.9000
Cc-b0027	34.5100	3.4100	9.7300	0.5480	2.5700	225.1000
Cc-b0028	33.4000	3.4800	8.2500	0.5570	2.5700	221.3000
Cc-b0029	35.7100	4.3000	8.3000	0.5470	2.1900	220.4000
Cc-b0030	35.2100	3.4400	8.8100	0.5550	2.3800	207.1000
Cc-b0031	31.1900	4.5800	5.8100	0.5500	2.8700	217.8000
Cc-b0032	31.0700	5.2100	5.2000	0.5680	2.8000	241.4000
Cc-b0033	30.6500	4.9100	5.2100	0.5510	2.5100	210.3000
Cc-b0034	31.7700	5.1100	4.7200	0.5490	2.7300	224.7000
Cc-b0035	31.0400	4.8500	4.3400	0.5440	3.3700	234.6000
Cc-b0036	32.2500	4.5800	4.9600	0.5460	2.9700	240.8000
Cc-b0037	32.1500	4.9800	3.3900	0.5620	2.4300	188.5000
Cc-b0038	31.6000	5.1400	5.3500	0.5700	2.4800	205.2000
Cc-b0039	31.8600	5.0000	4.4900	0.5370	3.3000	234.9000
Cc-b0040	31.0500	5.1800	5.1500	0.5410	2.7900	214.4000
Cc-b0041	30.7200	4.8200	5.0200	0.5390	2.8700	217.6000
Cc-b0042	33.1000	5.5000	4.5000	0.5360	2.8500	209.3000
Cc-b0043	30.9700	4.5800	5.4000	0.5360	2.6000	222.3000
Cc-b0044	30.6400	4.9000	5.6800	0.5370	2.7800	220.4000
Cc-b0045	31.1700	4.8700	4.8600	0.5510	2.6000	224.1000
Cc-b0046	32.4900	5.2900	4.5400	0.5530	2.5500	223.8000
Cc-b0047	31.6600	5.5300	4.7400	0.5410	2.6200	245.3000

Cc-b0048	31.0700	5.1500	4.7000	0.5670	2.4900	250.2000
Cc-b0049	31.0200	4.8000	5.1200	0.5500	2.5700	223.5000
Cc-b0050	31.7500	5.1900	4.1200	0.5550	2.7000	239.8000
Cc-b0051	32.4800	5.3600	3.3700	0.5420	2.4900	199.8000
Cc-b0052	33.1500	5.6700	2.9400	0.5600	2.5800	196.2000
Cc-b0053	32.4000	5.5600	4.6800	0.6070	2.3300	185.0000
Cc-b0054	31.6900	5.4900	5.7200	0.5440	2.3400	196.0000
Cc-b0055	31.1400	4.5200	5.1100	0.5380	2.4900	188.2000
56	33.3200	5.2800	5.2200	0.6000	2.3700	234.3000
57	32.9500	5.0000	4.7200	0.5580	2.3100	213.4000
58	32.5100	4.8600	4.8400	0.5770	2.2300	198.0000
59	31.1400	4.7500	4.9300	0.5610	2.5400	215.4000
60	33.1700	4.5000	6.0000	0.5850	2.2500	218.5000
62	31.4300	5.8000	4.0000	0.5280	2.8790	246.6000
63	31.9500	6.2200	3.5800	0.5295	2.8440	216.2000
64	30.7500	6.1600	4.0800	0.5265	2.6300	239.7000
65	33.5100	6.6400	3.7300	0.4795	2.3600	181.2000
66	30.1200	5.1800	4.4100	0.4885	3.1300	241.8000
67	31.5000	5.2500	4.8700	0.5015	2.4000	189.4000
68	31.7400	6.0000	2.9800	0.4815	2.4000	148.1000
69	31.4000	5.7600	3.0300	0.4925	2.0300	148.5000
70	31.5300	5.9100	3.6300	0.5160	1.9600	165.4000
71	32.3700	5.9300	4.0100	0.5255	2.1400	160.8000
72	30.6400	5.6400	2.9700	0.5115	2.6500	218.4000
73	30.6500	4.9200	6.1200	0.5280	2.4600	188.6000
74	30.3100	5.4800	5.5100	0.5360	1.8400	199.9000
75	31.5500	5.0900	5.1700	0.5350	2.4600	240.1000
76	31.6600	5.8700	5.1100	0.5340	2.5100	206.6000
77	31.4200	5.6900	3.6700	0.5485	1.8200	216.3000
78	31.3100	5.5400	4.5300	0.5515	2.2700	222.4000
79	32.3100	5.6300	3.1900	0.5330	2.1900	230.8000
80	33.1800	5.8300	4.8500	0.5410	2.2600	202.7000
81	31.8000	5.0800	5.1000	0.5300	2.2700	218.2000
82	32.5700	6.0200	4.5800	0.5175	2.1900	221.0000
83	32.2800	5.8300	3.8300	0.5260	2.2600	213.3000
84	29.7800	5.2500	4.3600	0.5175	2.4200	170.0000
85	31.0200	5.4300	5.1800	0.5205	2.6200	220.4000
86	30.1600	5.4500	3.9900	0.5335	2.4600	224.4000
87	31.6000	5.8300	3.8800	0.5930	2.4000	236.9000
88	28.4100	4.7100	3.5500	0.5105	2.0700	205.7000
89	30.4400	5.3700	5.2100	0.5385	2.1000	208.7000
90	29.5500	5.1100	3.9700	0.5265	2.2500	180.0000
91	29.3100	4.6300	4.1600	0.5206	2.8900	260.3000
92	29.8600	5.6400	4.7000	0.5515	3.0900	258.2000
93	31.8200	5.8300	6.3900	0.5200	1.8200	195.5000
94	28.9200	5.0600	4.2900	0.5255	2.8600	237.9000

### Appendix 3: season three laboratory data

Sample	L	a	b	wa	Nitrogen	TA
1	33.1700	5.8800	5.5500	0.5420	2.6820	240.1300
2	32.1900	5.4200	4.0200	0.5700	2.7100	205.4300
3	31.8000	5.4700	3.9800	0.5430	2.5750	189.8400
4	32.7500	5.3400	3.8300	0.5640	2.2140	185.8000
5	29.5400	4.6100	4.5600	0.5370	2.2910	204.9000
6	29.8200	4.7300	4.1800	0.5605	2.8780	176.6000
7	31.4000	5.7900	3.0300	0.5160	2.0430	178.4000
8	33.8800	5.8900	4.8300	0.5495	2.3760	200.9000
9	32.1000	4.9600	3.8100	0.5680	2.4920	185.3000
10	31.6900	5.7300	4.2100	0.5680	2.6040	202.0000
11	31.5800	5.1500	4.2200	0.5055	2.4260	202.4000
12	33.0400	5.7000	7.5400	0.5400	1.8485	224.6000
13	31.3700	5.6100	5.9100	0.5705	2.0860	182.1000
14	32.4200	5.8300	4.5300	0.5640	2.3800	204.7000
15	31.9000	5.3700	3.5900	0.5460	2.2630	184.0000
16	30.1400	5.5200	4.7300	0.5865	2.5390	176.8000
17	31.0000	5.6600	4.6800	0.5615	2.6550	200.9000
18	30.4500	4.7500	4.3800	0.5435	2.3270	220.6000
19	32.7900	5.4800	5.9500	0.5555	2.9840	235.7000
20	32.1800	5.1100	5.2500	0.5460	2.5510	189.7000
21	32.4900	4.9900	5.8100	0.5390	2.2620	185.6000
22	30.7300	5.1900	3.4700	0.5690	2.1800	184.0000
24	31.2200	4.9100	3.4800	0.5510	2.8420	203.7000
25	32.5500	4.8400	7.2100	0.5315	2.3270	191.7000
26	32.0800	5.5800	7.0100	0.5355	1.4760	259.0000
27	30.0600	4.1300	4.7600	0.5390	2.2770	233.3000
28	30.9400	5.1100	7.9800	0.5240	2.3410	233.1000
29	33.0900	5.6900	6.8200	0.5435	2.7180	198.0000
30	32.6400	5.7500	7.6600	0.5100	1.9300	216.9000
31	33.4300	6.2500	6.7900	0.5255	1.8290	212.8000
32	35.0600	6.0900	6.7000	0.5610	2.2400	276.8000
33	32.7100	5.4500	5.0000	0.5605	2.5250	218.5000
34	32.9100	5.5900	4.1700	0.5370	2.5860	200.0000
35	32.5400	5.2600	4.0400	0.5570	2.3630	191.8000
36	33.7800	5.5800	6.2700	0.5620	1.8460	214.5000
37	33.0400	5.8600	8.7700	0.5340	2.1140	209.8000
38	32.5600	5.5600	8.9500	0.5305	2.3080	260.6000
39	28.7600	4.7200	3.0300	0.5395	2.8590	359.3000
40	29.0900	4.8200	3.7500	0.5220	2.4160	258.1000
41	31.0000	5.1000	3.8900	0.5215	2.3970	210.3000
42	31.8600	5.3600	3.6000	0.5355	2.4540	184.9000
43	32.0600	5.2400	4.5700	0.5310	2.4030	219.4000
44	32.7700	5.8300	4.2400	0.5920	2.1250	199.7000
45	33.0400	5.4900	6.5600	0.5570	1.7240	223.2000
46	32.5800	6.1900	6.9800	0.5645	1.9250	205.3000
47	31.3300	5.1600	3.4800	0.5340	2.4860	199.3000
48	31.6600	5.0400	3.6700	0.5425	2.3610	228.4000

49	33.4400	5.5700	5.9400	0.5670	2.8900	272.2000
50	32.6500	5.5500	3.8500	0.5550	2.5250	209.8000
51	32.8300	5.1900	8.0000	0.5050	2.4090	218.6000
53	31.7900	5.4800	6.2500	0.5380	2.5480	238.6000
52	30.8700	5.3800	4.8200	0.5390	2.3720	215.0000
54	30.8500	5.4800	6.3900	0.5485	2.2120	242.5000
55	31.1500	5.2300	4.4400	0.5715	2.5060	199.2000
56	30.0500	5.1600	3.2200	0.5390	2.5900	197.7000
57	31.2100	5.4300	3.9000	0.5255	2.3610	189.4000
58	30.6800	6.0600	4.7000	0.5365	1.8870	172.6000
59	32.6000	4.9400	5.8200	0.5735	2.8900	223.1000
60	32.8500	4.9300	5.0300	0.5475	2.7910	177.5000
61	33.9300	5.0100	8.0500	0.5460	2.5250	168.3000
62	31.3900	6.0500	5.5200	0.5080	2.0890	177.5000
63	31.3800	6.1300	4.0800	0.5390	2.0340	212.9000
64	32.1600	6.5800	4.6200	0.5410	1.7470	181.7000
65	31.8400	6.7700	4.4100	0.5315	2.1030	202.0000
66	30.6200	6.1000	4.8300	0.5885	2.3990	209.5000
67	30.8200	6.2200	5.4800	0.5350	1.9280	192.2000
68	30.4800	5.6500	5.2800	0.5545	2.3980	200.6000
69	32.7400	6.2000	4.9300	0.5340	2.4390	237.3000
70	31.0100	6.1500	4.8900	0.6005	2.0360	185.9000
71	31.3300	5.7400	4.8900	0.5235	2.2240	212.6000
72	32.8500	6.4100	5.1100	0.5475	2.4060	216.8000
73	31.9100	5.9000	5.1200	0.5385	2.6520	250.2000
74	30.8500	5.5800	4.1300	0.5250	2.1910	186.3000
75	31.3000	5.1900	3.9100	0.5765	2.3050	238.8000
76	30.8400	4.4400	5.5800	0.5295	2.2360	192.1000
77	31.2100	4.5700	4.3100	0.5185	2.6180	186.0000
78	32.0100	5.2200	4.9200	0.5080	2.7760	208.2000
79	30.4600	4.9100	6.8900	0.5010	3.1240	215.3000
80	30.4300	4.9400	5.6500	0.4945	3.1730	219.3000
81	29.8400	5.2400	5.6600	0.5655	2.3950	209.2000
82	30.4700	4.7700	4.2200	0.5180	2.5440	197.4000
83	32.8900	5.3800	5.1000	0.5575	2.1480	185.1000
84	30.0100	5.1900	4.7200	0.5815	2.0150	206.7000
85	30.8000	5.1900	4.9500	0.5700	2.2520	187.0000
86	31.3200	5.8700	3.9300	0.5705	2.2190	213.9000
87	31.0400	5.7400	4.4500	0.5925	2.3890	217.5000
88	32.7000	5.6700	5.7900	0.5175	2.5440	247.5000
89	33.2300	4.7600	6.4400	0.5625	2.5720	204.4000
90	33.4300	4.7700	7.5600	0.5370	2.1720	192.2000
91	33.3100	4.9700	6.9400	0.5605	2.4000	193.5000
92	32.3600	4.5100	7.7100	0.5290	2.4700	195.0000
93	32.5900	4.5100	7.8300	0.5355	2.4810	203.5000
94	33.8000	5.2300	5.4800	0.5585	2.3560	200.3000
95	32.7100	5.5500	7.9600	0.5315	2.8650	181.3000
96	30.0700	4.1700	5.7000	0.5460	2.1720	275.5000
97	34.6800	4.0400	7.0400	0.5340	2.9400	205.1000
98	35.1400	5.5400	6.4300	0.6120	2.6660	209.4000
99	30.9400	5.1000	7.1900	0.5120	2.0743	229.3000

100	33.1300	4.1700	6.9800	0.5585	2.5260	187.7000
101	32.0600	5.2700	6.7200	0.5115	2.2380	248.3000
102	34.5100	5.2700	7.3500	0.5795	2.5800	219.2000
103	34.7300	3.4300	8.1300	0.5360	3.2050	216.5000
104	29.3700	4.2700	7.1300	0.5370	2.0190	228.5000
105	33.6100	5.7400	5.2900	0.5580	2.0460	256.7000
106	32.7900	3.3300	9.3500	0.5355	3.1360	215.6000
107	31.9800	4.5500	7.1800	0.5220	2.3780	242.1000
108	32.7900	3.0600	8.4600	0.5355	2.8830	210.5000
109	33.3800	4.4800	5.3300	0.5415	2.3600	230.1000
Cc-d0001				0.6055	2.1970	191.3000
Cc-d0002				0.5860	2.4050	200.5000
Cc-d0003				0.5880	2.2230	179.7000
Cc-d0004				0.5545	2.5020	257.6000
d-5				0.6075	2.1820	200.4000
d-6				0.5980	1.8850	193.3000
d-7				0.5935	2.1790	192.8000
d8				0.5625	2.4400	218.2000
d-9				0.5975	2.8260	231.0000
d-10				0.5855	2.0330	198.9000
d11				0.5620	2.0500	228.4000
d-12				0.5565	2.8630	223.9000
d-13				0.5610	2.3390	185.7000
d-14				0.5625	2.5800	227.4000
d-15				0.5445	2.3950	195.8000
d-16				0.5790	2.8430	207.6000
d-18				0.5540	2.5490	216.6000
d-17				0.5735	2.2220	212.1000
d-19				0.5685	1.6980	231.6000
d-20				0.5900	1.9920	195.0000
d-21				0.5595	2.4970	204.3000
d-22				0.6050	2.1180	193.6000
d-23				0.5620	2.3900	224.5000
d-24				0.6200	1.9800	199.7000
d-25				0.5555	2.1290	167.5000
d-26				0.5905	2.5000	193.6000
d-27				0.5880	2.0060	183.9000
d-28				0.6145	2.3460	192.5000
d-29				0.5705	2.3140	215.4000
d-30				0.5655	2.2860	212.6000
d-31				0.6010	2.2650	193.8700
d-32				0.5795	2.2580	193.9000
d-33				0.5570	1.9410	183.6000
d-34				0.6000	2.5580	255.8000
d-35				0.5940	2.6950	256.9000
d-36				0.5950	2.1800	205.0000
d-37				0.5790	2.1430	180.6000
d-38				0.5410	2.5640	210.4000
d-39				0.5520	2.3520	221.6000
d40				0.5810	1.9750	222.6000
d-41				0.5335	2.1450	254.6000

d-42	0.5515	2.6610	220.6000
d-43	0.5395	2.1230	247.8000
d-44	0.5625	2.9100	278.8000
d-45	0.5710	1.8980	231.6000
d-46	0.5990	1.8000	203.9000
d-47	0.5930	2.5540	236.6000
d-48	0.6000	2.0450	201.5000

## Appendix 4: laboratory data for airstream sorting

Sample	L*	a*	b*	Water Activity	TA
as-1	37.2400	4.4400	5.9800	0.4660	315.7795
as-2	35.0300	4.5600	9.1300	0.4240	208.7698
as-3	34.8000	4.2900	9.0700	0.4185	202.4974
as-4	29.1300	3.1500	8.0400	0.4450	204.2560
as-5	29.3000	3.0500	9.0600	0.4480	216.5701
as-6	33.3400	4.2500	4.9100	0.4595	195.6512
as-7	38.4500	4.8800	7.4400	0.4655	174.1802
as-8	33.7200	4.0500	10.6600	0.4160	210.5279
as-9	33.9100	3.9800	11.5900	0.4185	210.0567
as-10	36.5900	4.6600	7.5800	0.4090	216.6842
as-11	33.5700	3.6000	10.8600	0.4050	227.6972
as-12	36.3800	4.9700	7.9600	0.4020	209.2774
as-13	36.7000	4.8600	8.0600	0.3915	201.9547
as-14	35.7300	5.1000	5.7200	0.4005	240.4066
as-15	35.9700	4.6500	8.0600	0.3995	201.3892
as-16	30.5300	2.8600	9.6500	0.4610	302.1028
as-17	29.6000	2.4200	9.2800	0.4550	277.5779
as-18	30.4200	2.4800	9.9800	0.4510	301.8352
as-19	29.9200	2.5300	9.6600	0.4535	293.9608
as-24	27.8100	3.3500	5.6800	0.4775	215.2247
as-25	28.4700	2.7700	5.6800	0.4510	210.7171
as-26	31.8400	3.5100	10.3700	0.4435	190.6878
as-27	33.4500	3.5900	8.2700	0.4455	188.7804
as-20	30.4700	2.5000	10.1400	0.4590	318.7841
as-21	31.6200	3.0300	9.7000	0.4560	353.3865
as-22	29.2300	2.3600	7.9400	0.4470	282.0731
as-23*	34.4600	3.2600	9.7000	0.4455	268.3903
as-28	31.0100	3.5100	8.2700	0.4335	198.6581
as-29	29.5800	3.0000	7.4500	0.4370	212.5915
as-30	32.0800	3.6900	10.5900	0.4575	193.0905
as-31	32.0000	3.2800	7.9900	0.4425	201.0584
as-32	34.4200	4.1100	10.7100	0.4420	212.8170
as-33	33.8000	3.6800	9.5000	0.4130	217.1912
as-34	34.1100	3.9300	11.0500	0.4120	221.3860
as-35	32.1500	3.8600	12.1700	0.4135	233.0042
as-36*	37.1400	4.8500	9.3400	0.4240	209.6757
as-37	37.3000	4.7200	7.1300	0.3970	219.5246
as-38	31.4100	3.8000	12.0900	0.3995	218.7424
as-39	31.9300	3.9100	11.4200	0.3990	228.1416
as-40	30.6400	3.9500	7.9000	0.4745	173.6474
as-41	37.5600	5.0200	6.2000	0.4560	159.2599
as-42	37.9700	5.2600	6.5300	0.4535	158.5366
as-43	38.5600	5.0900	5.7900	0.4535	171.2111
as-44	37.7600	5.1300	5.2400	0.4515	177.0985
as-45	38.9400	5.1000	5.8600	0.4490	161.2863
as-46	37.7900	4.8500	6.4500	0.4630	164.0540
as-47	39.9700	5.4200	6.7300	0.4610	152.9512

as-48	32.7400	4.0400	3.6000	0.4375	231.2526
as-49	32.4400	3.9000	2.0600	0.4390	216.6364
as-50	35.4600	4.3600	3.8800	0.4420	213.4651
as-51	33.2900	3.7900	6.0400	0.4390	213.5409
as-52	32.7500	3.9500	5.7800	0.4565	210.2420
as-53	30.0100	2.8400	6.0300	0.4435	210.2729
as-54	29.0200	2.8200	7.3200	0.4405	218.8933
as-55	28.3100	2.7100	6.9700	0.4395	217.8441
as-56	30.8300	2.9800	6.2100	0.4785	191.8738
as-57	34.6400	4.3700	2.2100	0.4660	196.2555
as-58	33.5800	4.3300	2.3200	0.4635	201.7068
as-59	30.8300	3.7900	2.9900	0.4635	207.3604
as-60	34.6300	3.9500	4.0800	0.4615	170.9377
as-61	34.9700	4.3700	2.4700	0.4660	200.7072
as-62	33.5500	3.6000	4.2700	0.4595	188.7116
as-63	33.3200	3.7100	5.0200	0.4615	171.9115
as-64	36.0400	4.3400	5.2500	0.4345	251.5494
as-65	33.0200	4.0800	10.1300	0.4640	236.8539
as-66	34.7100	4.6700	8.0200	0.4290	275.0169
as-67	33.2700	4.1300	7.1300	0.4895	277.4840
as-68	34.0500	4.7100	6.1100	0.5315	223.7190
as-69	34.0300	4.5700	8.8900	0.4650	223.0217
as-70	34.6600	4.0400	8.6200	0.4660	313.1748
as-71	34.2200	4.3400	9.0800	0.4615	281.8182
as-72	35.2900	4.0900	8.8900	0.5060	217.7375
as-73	35.1800	4.1800	7.2100	0.4740	255.2203
as-74	39.4000	4.6000	8.5000	0.4825	214.1796
as-75	36.8600	5.9400	6.6800	0.4925	286.2794
as-76	36.6000	4.1800	6.8900	0.4610	291.5708
as-77	34.4600	4.9200	6.9800	0.5145	290.6902
as-78	36.6100	4.0200	6.5500	0.4550	221.0445
as-79	35.2400	4.0100	5.1100	0.4800	235.2556
as-80	36.7900	5.1300	4.7400	0.5085	228.6109
as-81	37.8900	5.0100	5.5000	0.4810	242.2547
as-82	38.5400	5.1600	6.6300	0.4840	265.2473
as-83	36.3800	4.9400	4.4000	0.5270	230.9534
as-84	37.7300	4.2900	6.6200	0.5015	278.3193
as-85	37.8500	4.5700	6.9200	0.4780	228.3099
as-86	37.0300	5.0100	6.0200	0.4745	251.7192
as-87	38.4800	4.6300	6.7600	0.4530	276.4296
as-88	39.2400	4.0700	7.8400	0.4925	241.1334
as-89	34.7100	4.4000	8.1200	0.4995	193.8631
as-90	32.3400	3.3500	10.4800	0.4610	259.9623
as-91	32.1800	3.6500	9.2600	0.4595	271.7187
as-92	34.0000	3.7500	10.1600	0.4870	248.4836
as-93	33.6800	3.4700	8.9800	0.4700	220.5978
as-94	34.0300	4.0200	8.9700	0.4545	207.1208
as-95	34.4600	3.3000	10.7000	0.4885	206.2924
CC-AS0096	32.9400	2.7100	9.3400	0.4425	316.5441
as-97	34.5300	3.9800	10.5300	0.4500	222.9927
as-98	29.8800	2.6500	9.9700	0.4305	256.2550

as-99	31.7300	3.2900	9.3300	0.4790	263.6635
as-100	31.5700	4.8700	9.1600	0.4875	193.9550
as-101	32.3600	4.1900	8.8300	0.4215	209.0129
as-102	35.0000	4.0700	8.5700	0.4595	225.8789
as-103	34.1000	5.0700	11.4200	0.4895	186.7571
as-104	32.2900	3.8000	8.0000	0.4730	228.7715
as-105	35.7900	4.4000	7.6700	0.5320	212.2502
as-106	35.0900	3.5100	5.8900	0.5055	261.0976
as-107	33.0500	5.0700	10.3100	0.4435	185.9332
as-108	33.1700	4.1300	8.6300	0.4625	205.7896
as-109	37.3000	5.0100	6.3000	0.4525	191.5405
as-110	35.7300	4.2700	6.6200	0.4840	275.4001
as-111	34.7100	3.6700	9.4900	0.4840	276.6725
as-113	33.8700	4.5300	9.9700	0.4870	207.4328
as-114	34.0100	3.9400	10.0500	0.4585	182.9826
as-115	35.9800	4.9200	6.1100	0.4935	193.5974
as-116	35.7100	3.5400	5.9300	0.4405	466.7969
as-117	35.8400	4.3000	8.8000	0.4620	238.2915
as-118	36.8600	5.6300	7.4100	0.4810	166.3582
as-119	35.8500	3.8100	10.0100	0.4630	224.6741
as-120	33.6400	4.1800	8.1000	0.4750	222.0264
as-121	35.7300	4.8200	7.7700	0.4920	174.9736
as-122	35.8300	4.3900	7.8100	0.4795	189.6155
as-123	35.4200	4.0500	9.5700	0.4280	211.2045
as-124	38.6800	5.0500	6.9000	0.4745	241.4165
as-125	38.3100	5.6100	6.8200	0.5130	156.3607
as-126	36.9600	4.4500	5.5900	0.4675	281.2170
as-127	37.6900	4.6400	5.3200	0.4840	237.9036
as-128	38.2200	5.0700	7.1000	0.4965	218.7955
as-129	36.8800	4.1100	7.2500	0.5005	219.2235
as-130	36.7000	4.3900	5.7800	0.4890	238.6460
as-131	38.8300	5.5500	7.1800	0.4960	159.1072
as-132	35.9200	4.5600	8.7400	0.4855	212.8749
as-133	36.1200	3.5700	7.8100	0.4535	264.6821
as-134	36.2900	5.0900	6.6400	0.4925	162.4318
as-135	37.9500	4.2300	9.6700	0.4640	248.1469

## **Appendix 5: Experimental procedures undertaken at CSIRO, Merbien for the sultana maturity by air-steam sorting**

*These samples were provided by Peter Clingeleffer et al at CSIRO, Merbien to develop an NIRS calibration to determine fruit maturity.*

### **5.0 Source of Samples**

Samples of fruit were all sourced from the CSIRO Merbein South vineyard, the vines were mature own rooted sultana vines (25+ years old) grown on commercial T-trellis and irrigated by under-vine sprinkler irrigation (previously drip irrigation).

### **5.1 Processing of samples**

Each week from the tenth of February 2004, 50kg samples of fruit were collected from several vines along the row. This fruit was sub-sampled for juice analysis (refer to table below) then the bunches were dipped in drying emulsion for 3 minutes, removed, drained and later spread on a drying rack. The emulsion used to treat the samples consisted of a mix which included 1% proprietary dipping oil and 1.25% Potash (potassium carbonate) in water. The drying racks used to dry the samples were about 50m long and had tiers of wire netting 'shelves' (8 shelves high) with a galvanised iron roof. The racks used are the equivalent of those used by the sultana industry. The bunches after dipping were drained and placed on the racks drying tiers to dry. Normally, it would take between 3 and 6 weeks to dry the fruit, depending on the weather. The first 3 harvest dates took 3 weeks to dry, the later ones required 5 weeks. The dried fruit was removed from the rack when it reached about 13% moisture, and were then finish dried in the sun. Finish drying entails spreading the dried grapes on a black plastic sheet out in the sun for 1-3 days. In this case due to good hot weather the finish drying was only 1 day and the final moisture content was 10-11%w/w. Once dried the fruit was stored at ambient temperature for a few months and then stored in a 2°C low humidity cool-room.

## 5.2 Data collected at CSIRO, Merbein of these samples

Below is a copy of the fruit characteristics measured on the sub-sample taken at harvest plus some dried measurements made prior to airstream sorting.

Date picked	Type	Position	°Brix	pH	Acid (g/l)	Fresh MBW g	Dried 100 berry wt (g)	L (*)	a (*)	b (*)	Moisture
10/02/2004	sultana	MA15 v4/5	21.3	3.36	7.68	1.45	33.89	25.94	4.06	12.39	11.3
17/02/2004	sultana	MA15 v6/9	21.9	3.61	5.68	1.42	34.64	28.40	3.80	13.63	10.5
24/02/2004	sultana	MA15 v9/12	23.1	3.67	5.6	1.39	32.31	29.65	5.07	14.43	10.5
2/03/2004	sultana	MA15 v12/15	23.9	3.72	5.04	1.41	32.32	22.46	4.98	9.41	11
9/03/2004	sultana	MA15 v16/18+22	24	3.73	6.24	1.76	34.83	24.09	6.22	9.80	11.3
16/03/2004	sultana	MA15 v22/25	24.3	3.78	4.95	1.83	42.21	25.67	5.93	11.51	11.3
23/03/2004	sultana	MA15 v25/28	24	3.83	4.52	1.60	40.66	30.76	4.63	15.94	11.5

\*Date readings taken, 14/7/04

See Chapter 6 for details of use of these provided samples.

Reference: C. Tarr, D. Emmanuelli, K. Connolly and P. Clingeleffer, CSIRO, Merbein, Victoria.: Air-stream Sorting Procedures.