

THESIS

**DETERMINATION OF CHEMICAL COMPOSITIONS AND
PULP YIELD OF *EUCALYPTUS CAMALDULENSIS* WOOD BY
NEAR INFRARED SPECTROSCOPY**

NATTAPORN SUTTIWIJITPUKDEE

GRADUATE SCHOOL, KASETSART UNIVERSITY

2006



THESIS APPROVAL
GRADUATE SCHOOL, KASETSART UNIVERSITY

Master of Science (Chemistry)

DEGREE

Industrial Chemistry

FIELD

Chemistry

DEPARTMENT

TITLE: Determination of Chemical Compositions and Pulp Yield of *Eucalyptus camaldulensis* Wood by Near Infrared Spectroscopy

NAME: Miss Nattaporn Suttiwijitpukdee

THIS THESIS HAS BEEN ACCEPTED BY

Vittaya Punsuvon

THESIS ADVISOR

(Associate Professor Vittaya punsuwon, Ph.D.)

Cholticha Noomhorm

COMMITTEE MEMBER

(Associate Professor Cholticha Noomhorm, Ph.D.)

Warunee Thanapase

COMMITTEE MEMBER

(Mrs. Warunee Thanapase, M.S.)

Atchana Wongchaisuwat

COMMITTEE MEMBER

(Assistant Professor Atchana Wongchaisuwat, M.S.)

Yerry Mahatumarattana

ACTING
DEPARTMENT HEAD

(Assistant Professor Yerry Mahatumarattana, B.Sc.)

APPROVED BY THE GRADUATE SCHOOL ON 05 April 2006

Vinai Artkongharn

DEAN

(Associate Professor Vinai Artkongharn, M.A.)

THESIS

DETERMINATION OF CHEMICAL COMPOSITIONS AND PULP YIELD OF *EUCALYPTUS CAMALDULENSIS* WOOD BY NEAR INFRARED SPECTROSCOPY

NATTAPORN SUTTIWIJITPUKDEE

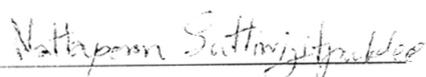
**A Thesis Submitted in Partial Fulfillment of
the Requirements for the Degree of
Master of Science (Chemistry)
Graduate School, Kasetsart University**

2006

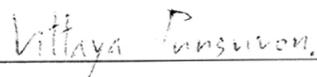
ISBN 974-16-1427-6

Nattaporn Suttiwijitpukdee 2006: Determination of Chemical Compositions and Pulp Yield of *Eucalyptus camaldulensis* Wood by Near Infrared Spectroscopy. Master of Science (Chemistry), Major Field: Industrial Chemistry, Department of Chemistry. Thesis Advisor: Associate Professor Vittaya Punsuvon, Ph.D. 160 pages.
ISBN 974-16-1427-6

Pulps and paper industry is one of the important industry in Thailand. It uses a lot of chemical reagents, and time consuming for analyzing the quality of Eucalyptus wood as the raw material. The purpose of this study is to determine the possibility replacing the conventional method by the near infrared spectroscopic (NIRs) method as a rapid method for the chemical quantitative of Eucalyptus wood mill. Seventy samples of Eucalyptus wood mill were measured by NIR in the region of 1100-2500 nm and the spectra were pretreated by Multiplicative Scatter Correction (MSC) and the second derivative statistic. The calibration model was developed by Multiple Linear Regression (MLR) and Partial Least Square Regression (PLSR) to predict total pulp yield, screen pulp yield, and the chemical properties such as the holocellulose, α -cellulose, pentosan, lignin, wood extractive, ash, 1% NaOH solubility, and monosaccharide contents. The statistic results of MLR model showed the correlation coefficient (R) in range 0.704-0.967 and the standard errors of prediction (SEP) in range of 0.063-1.670%. While the PLS model performed R value in range of 0.720-0.970 and SEP values in the range of 0.061-1.496%. The results were shown that NIR technique has a high potential for the evaluation of milled wood properties. Especially, MLR model for determination of total pulp yield provided R and SEP of 0.964 and 0.945, respectively. MLR model for determination of screen pulp yield showed R and SEP of 0.967 and 0.843, respectively. MLR models of 1%NaOH solubility found R and SEP of 0.945, and 0.702, respectively. All of those models provided values of RPD over 3.



Student's signature



Thesis Advisor's signature

23 / 03 / 06

ACKNOWLEDGEMENTS

I'm most grateful to thank and recognize Associate Professor Vittaya Punsuvon, who is the chairman of my graduate committee for the invaluable suggestion, approval, correction, and overall through my thesis.

Appreciation is also extended to others member of the thesis committee, Associate Professor Cholticha Noomhorm and Assistant Professor Atchana Wongchaisuwat for teaching and their helpful comments.

I am grateful to The Siam Pulp and Paper Public Co., Ltd. for the sample and some information support of this research.

Moreover, I'm also sincerely grateful to Mrs.Warunee Thanapase at Kasetsart Agricultural and Agro-industrial Product Improvement Institute (KAPI) for her encouragement, very useful suggestions and support.

My special thanks are extended to Dr. Sumaporn Kasemsumran. Without her knowledge in programming and analyze, this project would never have been completed.

Thanks to Nanthaya, I really appreciate your help with the experiments. I'm also extremely grateful to my entire friend for cheerful, helps, and stand beside me during my life in Kasetsart University.

Finally, I thank my parents and my family for their support through out my education.

Nattaporn Suttiwijitpukdee

February 2006

TABLE OF CONTENTS

| | Page |
|---|-------------|
| TABLE OF CONTENTS | i |
| LIST OF TABLES | ii |
| LIST OF FIGURES | vi |
| LIST OF ABBREVIATIONS | xiv |
| INTRODUCTION | 1 |
| LITERRATURE REVIEWS | 2 |
| General Background | 2 |
| The Chemical Composition of <i>Eucalyptus camaldulensis</i> | 2 |
| Pulping | 6 |
| Description of Chemical Method in Pulp and Paper Industry | 11 |
| Principles of Near Infrared Spectroscopy | 12 |
| Sample Presentation | 18 |
| Chemometric Methods for Quantitative Analysis | 19 |
| Application of Near Infrared Spectroscopy for Wood Industry | 27 |
| MATERIALS AND METHODS | 32 |
| Materials | 32 |
| Equipments | 32 |
| Methods | 33 |
| RESULTS AND DISCUSSION | 40 |
| Raw Material | 40 |
| Chemical Analysis | 40 |
| NIR Analysis | 45 |
| The Data Analysis | 58 |
| CONCLUSION | 124 |
| LITERATURE CITED | 125 |
| APPENDIX | 130 |
| Appendix A | 131 |
| Appendix B | 158 |

LIST OF TABLES

| Table | | Page |
|--------------|---|-------------|
| 1 | Divisions of the infrared region | 13 |
| 2 | Summary of statistics for the measurements of pulp yield property | 40 |
| 3 | Summary of the chemical composition of wood powder | 41 |
| 4 | Summary of the carbohydrate component value of wood powder | 42 |
| 5 | Present recovery of glucose that received in pure cellulose | 43 |
| 6 | Tentative assignment of NIR bands for <i>E.calmaldulensis</i> woods mill and its components (Burn,2001, Ozaki,2001Osborn,1997, Shimleck,1997) | 51 |
| 7 | Tentative assignment of NIR bands for monosaccharide standards (Burn,2001, Ozaki,2001,Osborn,1997, Shimleck,1997). | 55 |
| 8 | Relationship of chemical component property, range of content, average of standard derivation of 70 samples. | 58 |
| 9 | Relation of sugar property, range of content, range of content and standard deviation of 70 samples | 59 |
| 10 | The statistical summaries of total pulp yield content calibration model with different pretreatment methods | 62 |
| 11 | The statistical summaries of total pulp yield content calibration model with vary three wavelength regions | 63 |
| 12 | The statistical summaries of screen pulp yield content calibration model with different pretreatment method. | 66 |
| 13 | The statistical summaries of screen pulp yield content calibration model with vary three wavelength regions | 67 |
| 14 | The statistical summaries of holocellulose calibration model with different present methods | 71 |
| 15 | The statistical summaries of holocellulose calibration model with vary three wavelength regions | 72 |
| 16 | The summary of statistic of α -cellulose calibration model that vary the pretreatment | 75 |

LIST OF TABLES (Cont'd)

| Table | | Page |
|--------------|--|-------------|
| 17 | The statistical summaries of α -cellulose calibration model that vary wavelength regions | 76 |
| 18 | The statistical summaries of pentosan calibration model with different pretreatment methods in vary method | 80 |
| 19 | The statistical summaries of pentosan calibration model that vary wavelength regions | 81 |
| 20 | The statistical summaries of lignin calibration model with different pretreatment methods. | 84 |
| 21 | The summaries of the statistic in lignin calibration model with vary two wavelength regions | 85 |
| 22 | The statistical summaries of 1% NaOH solubility value calibration model with pretreatment in vary method | 88 |
| 23 | Table The statistical summaries of 1% NaOH solubility calibration model with vary wavelengths regions | 89 |
| 24 | The statistical summaries of ash content calibration model with different pretreatment methods | 92 |
| 25 | The statistical summaries in ash content calibration model that vary wavelengths regions | 93 |
| 26 | The statistical summaries in wood extractives calibration model with different pretreatment methods | 96 |
| 27 | The statistical summaries in wood extractives calibration model with vary wavelength regions | 97 |
| 28 | The statistical summaries in glucose content calibration model with pretreatment in vary method | 100 |
| 29 | The statistical summaries in glucose content calibration model that vary wavelength regions | 101 |
| 30 | The statistical summaries in xylose content calibration model with different pretreatment methods | 104 |

LIST OF TABLES (Cont'd)

| Table | | Page |
|-----------------------|--|-------------|
| 31 | The statistical summaries in xylose content calibration model with vary wavelength regions | 105 |
| 32 | The statistical summaries in arabinose content calibration model with different pretreatment methods | 108 |
| 33 | The statistical summaries in arabinose content calibration model that vary wavelength regions | 109 |
| 34 | The statistical summaries in galactose content calibration model with different pretreatment methods | 112 |
| 35 | The statistical summaries in galactose content calibration model that vary wavelength regions | 113 |
| 36 | The statistical summaries in mannose content calibration model with pretreatment in vary method | 116 |
| 37 | The statistical summaries of mannose content calibration model that vary wavelength regions | 117 |
| 38 | The summary of statistic results in all components | 121 |
| | | |
| Appendix Table | | |
| A1 | The statistic result of total pulp yield calibration | 132 |
| A2 | The statistic result of the screen pulp yield calibration | 133 |
| A3 | The statistic result of the holocellulose calibration | 134 |
| A4 | The statistic result of the α -cellulose calibration | 136 |
| A5 | The statistic result of the pentosan calibration | 137 |
| A6 | The statistic result of the lignin content calibration | 138 |
| A7 | The statistic result of the 1%NaOH solubility value calibration | 140 |
| A8 | The statistic result of ash content calibration | 143 |
| A9 | The statistic result of wood extractives calibration | 146 |
| A10 | The statistic result of glucose content calibration | 149 |

LIST OF TABLES (Cont'd)

| Appendix Table | | Page |
|-----------------------|---|-------------|
| A11 | The statistic result of xylose content calibration | 151 |
| A12 | The statistic result of arabinose content calibration | 152 |
| A13 | The statistic result of galactose content calibration | 154 |
| A14 | The statistic result of mannose content calibration | 156 |

LIST OF FIGURES

| Figure | | Page |
|--------|--|------|
| 1 | Partial structure of cellulose | 2 |
| 2 | Abbreviated formula of glucuronoxylan | 3 |
| 3 | Abbreviated formula of glucomannan | 3 |
| 4 | The structure of three precursors of the lignin | 4 |
| 5 | The partial structure of lignin | 5 |
| 6 | The alditol acetate derivatisation reaction of xylose | 6 |
| 7 | Cleavage of β -Aryl Ether Bonds in nonphenolic phenylpropane units during soda pulping | 7 |
| 8 | Main reactions of phenolic β -aryl ether structures during alkali (soda) and kraft pulping. R=H, alkyl, or aryl group | 8 |
| 9 | Elimination of proton and formaldehyde from the quinone methide in intermediate during alkali pulping | 8 |
| 10 | Cleavage of α -aryl ether bonds in free phenolic structures | 9 |
| 11 | Cleavage of methyl aryl ether bonds with simultaneous formation of methyl mercaptan (CH_3SH), dimethyl sulfide (CH_3SCH_3), and dimethyl disulfide (CH_3SSCH_3) during kraft pulping | 9 |
| 12 | Examples of condensations reactions during alkali and kraft pulping | 10 |
| 13 | Peeling and stopping reactions of polysaccharides | 11 |
| 14 | Energy levels of a molecule. E_1, E_2 : electronic energy levels. $\nu' = 0, 1, 2, \dots, \nu' = 0, 1, 2, \dots$: vibration energy levels. $J'' = 0, 1, 2, \dots, J' = 0, 1, 2, \dots$ rotational energy levels | 13 |
| 15 | Present various region of spectroscopy | 15 |
| 16 | Diagram of NIR analysis concept | 16 |
| 17 | NIR spectrophotometer: Schematic diagram of a dispersive scanning grating near in-infrared spectrophotometer | 17 |
| 18 | Optical Head and Monochromator | 18 |
| 19 | Near Infrared Spectroscopic Instrument of Infra-Alyzer 500(Bran+ Luebbe) | 18 |

LIST OF FIGURES (Cont'd)

| Figure | | Page |
|---------------|--|-------------|
| 20 | Sample presentations of transmission(a), reflection(b), tranflection (c) and interaction(d) | 19 |
| 21 | Forty corn calibration samples: (A) untreated and (B) MSC applied to 40 calibration samples | 20 |
| 22 | Original Spectrum that combine from two Gaussian curves (a and b) compared with the Second derivative spectrum | 21 |
| 23 | Scheme for explanation of MWPLSR | 24 |
| 24 | The concept of PLS calculation. | 26 |
| 25 | Typical absorbance spectra of paper, showing absorbance regions attributed to cellulose and water | 30 |
| 26 | Presented the method of control moisture content in wood powder by drying in oven at 50°C. | 33 |
| 27 | Closed cup packing with wood (a) and sample holder using in the near infrared spectrometer InfraAlyzer 500 (Bran+Luebbe) (b). | 34 |
| 28 | <i>Eucalyptus camaldulensis</i> wood powder | 40 |
| 29 | Relationship between Total pulp yield and screen pulp yield | 41 |
| 30 | GC chromatogrm of alditol acetate derivative of monosaccharide in <i>Eucalyptus</i> wood powder | 42 |
| 31 | The correlations between the screen pulp yield value and α -cellulose content (a), the screen pulp yield and glucose content (b), α -cellulose content and glucose content (c) | 44 |
| 32 | An absorbance NIR and its second derivative spectrum of <i>Eucalyptus camaldulensis</i> wood powder | 45 |
| 33 | Seventy original near infrared spectra of <i>Eucalyptus camaldulensis</i> wood powder. | 46 |
| 34 | Averaged raw near infrared spectra of 70 <i>Eucalyptus camaldulensis</i> wood powder | 46 |

LIST OF FIGURES (Cont'd)

| Figure | | Page |
|---------------|---|-------------|
| 35 | MSC pretreated spectra of 70 <i>Eucalyptus camaldulensis</i> wood powder samples | 47 |
| 36 | Savitzky-Golay 2 nd derivative spectra of 70 <i>Eucalyptus camaldulensis</i> wood powder. | 48 |
| 37 | Second derivative spectra of <i>Eucalyptus camaldulensis</i> wood powder, holocellulose, and α -cellulose content | 49 |
| 38 | Second derivative spectra of lignin, wood extractive, ash from wood powder. | 50 |
| 39 | Second derivative spectra of glucose, xylose, arabinose standard | 53 |
| 40 | Second derivative spectra of galactose, mannose standard | 54 |
| 41 | Second derivative multiple correlation coefficients of total pulp yield for select the first wavelength in set up calibration model | 60 |
| 42 | The correlation plots of total pulp yield content between predicted value and the actual value of calibration set (a) and validation set (b) | 61 |
| 43 | The residue lines obtained by MWPLSR for total pulp yield | 62 |
| 44 | The regression coefficients from equation 22 of total pulp yield | 64 |
| 45 | The correlation plots of total pulp yield content between predicted value and the measure value of calibration set (a) and validation set (b) | 64 |
| 46 | The multiple correlation coefficients of screen pulp yield that pretreated with second derivative | 65 |
| 47 | The correlation plots of screen pulp yield between predicted value and actual value of calibration set (a) and validation set (b) | 66 |
| 48 | The residue lines obtained by MWPLSR for screen pulp yield | 67 |
| 49 | The regression coefficients from equation 24 with using wavelength in the region 1100-1250nm and 2030-2400 nm | 68 |

LIST OF FIGURES (Cont'd)

| Figure | | Page |
|---------------|---|-------------|
| 50 | The correlation plots of screen pulp yield between predicted value and the measured value of calibration set (a) and validation set (b) | 69 |
| 51 | The multiple correlation coefficients of holocellulose that pretreated with second derivative. | 69 |
| 52 | The correlation plots of holocellulose between predicted value and the actual value of calibration set (a) and validation set (b) | 70 |
| 53 | The residual lines obtained by MWPLSR for holocellulose | 71 |
| 54 | The regression coefficients from equation 26 with use in region 2180-2280 nm | 73 |
| 55 | The correlation plots of holocellulose between predicted value and measured value of calibration set (a) and validation set (b) | 73 |
| 56 | The multiple correlation coefficients of α -cellulose that pretreated with second derivative. | 74 |
| 57 | The correlation plots of α -cellulose between predicted value and the measure value of calibration set (a) and validation set (b) | 75 |
| 58 | The residue lines obtained by MWPLSR for α -cellulose | 76 |
| 59 | The regression coefficient from equation 28 with using wavelength in the region 1100-2500 nm | 77 |
| 60 | The correlation plots of α -cellulose between predicted value and the measured value of calibration set (a) and validation set (b) | 78 |
| 61 | The multiple correlation coefficients of pentosan that pretreated with second derivative | 78 |
| 62 | The correlation plots of pentosan between predicted value and the actual value of calibration set (a) and validation set (b) | 79 |
| 63 | The residue lines obtained by MWPLSR for Pentosan | 80 |

LIST OF FIGURES (Cont'd)

| Figure | | Page |
|---------------|---|-------------|
| 64 | The correlation coefficients of equation 30 with using wavelength in the region of 1100-2500 nm | 81 |
| 65 | The correlation plots of pentosan between predicted value and the measured value of calibration set (a) and validation set (b). | 82 |
| 66 | The multiple correlation coefficients of lignin that pretreated with second derivative | 82 |
| 67 | The correlation plot between predicted value and the measured value of calibration set (a) and validation set (b) | 83 |
| 68 | The residue lines obtained by MWPLSR for lignin | 84 |
| 69 | The regression coefficients from equation 32 with using wavelength in the region of 1480-1740 nm and 2030-2500 nm. | 85 |
| 70 | The correlation plots of lignin between predicted value and the measure value of calibration set (a) and validation set (b) | 86 |
| 71 | The multiple correlations of 1% NaOH solubility that pretreated with second derivative | 86 |
| 72 | Scatter plot between predicted value and the measure value of calibration set (a) and validation set (b) | 87 |
| 73 | The residue lines obtained by MWPLSR for 1% NaOH solubility | 88 |
| 74 | The regression coefficients from equation 34 of 1%NaOH solubility | 90 |
| 75 | The correlation plots of 1%NaOH solubility between predicted value and actual value of calibration set(a) and validation set(b) | 90 |
| 76 | The multiple correlation coefficients of ash content that pretreated with second derivative. | 91 |
| 77 | The correlation plots of ash content between predicted value and measured value of calibration set (a) and validation set (b) | 92 |
| 78 | The residue lines obtained by MWPLSR for ash content | 93 |
| 79 | The regression coefficient from equation 36 of ash with using wavelength 1100-1400 nm and 1880-2080 nm | 94 |

LIST OF FIGURES (Cont'd)

| Figure | | Page |
|---------------|--|-------------|
| 80 | The correlation plots of ash content between predicted value and actual value of calibration set (a) and validation set (b) | 94 |
| 81 | The multiple correlation coefficients of extractives contents that pretreated with second derivative | 95 |
| 82 | The correlation plots extractives content between predicted value and actual value of calibration set (a) and validation set (b) | 96 |
| 83 | The residue lines obtained by MWPLSR for extractive content | 97 |
| 84 | The regression coefficients from equation 38 of extractive with using wavelength 1100-1970 nm and 2090-2420 nm. | 98 |
| 85 | The correlation plots of extractive content between predicted value and actual value of calibration set(a) and validation set(b) | 98 |
| 86 | The multiple correlation coefficients of glucose content that pretreated with second derivative. | 99 |
| 87 | The correlation plots of glucose content between predicted value and actual value of calibration set (a) and validation set (b) | 100 |
| 88 | The residue lines obtained by MWPLSR for glucose content | 101 |
| 89 | The regression coefficients from equation 40 of glucose with using wavelength 1430-1840 nm | 102 |
| 90 | The correlation plots of glucose content between predicted value and actual value of calibration set (a) and validation set (b). | 102 |
| 91 | The multiple correlation coefficients of xylose that pretreated with second derivative | 103 |
| 92 | The correlation plots of xylose content between predicted value and actual value of calibration set (a) and validation set (b). | 104 |
| 93 | The residue lines obtained by MWPLSR for xylose | 105 |
| 94 | The regression coefficients of xylose from equation 42 with using wavelength in the region 1200-1280 nm and 1740-2090 nm | 106 |

LIST OF FIGURES (Cont'd)

| Figure | | Page |
|---------------|---|-------------|
| 95 | The correlation plots of xylose content between predicted value and actual value of calibration set (a) and validation set (b) | 106 |
| 96 | The multiple correlation coefficients of arabinose that pretreated with second derivative. | 107 |
| 97 | The correlation plots of arabinose content between predicted value and actual value of calibration set (a) and validation set (b) | 108 |
| 98 | The residue lines obtained by MWPLSR for arabinose | 109 |
| 99 | The regression coefficients of arabinose from equation 44 with using wavelength in the region of 1100-2500 nm | 110 |
| 100 | The correlation plots of arabinose content between predicted value and actual value of calibration set(a) and validation set(b) | 110 |
| 101 | The multiple correlation coefficients of galactose that pretreated with second derivative. | 111 |
| 102 | The correlation plots of galactose content between predicted value and measured value of calibration set (a) and validation set (b) | 112 |
| 103 | The residue lines obtained by MWPLSR for galactose | 113 |
| 104 | The regression coefficients of galactose from equation 46 with using wavelength in the region 1100-2500 nm | 114 |
| 105 | The correlation plots of galactose content between predicted value and measure value of calibration set (a) and validation set (b) | 114 |
| 106 | The multiple correlation coefficients of mannose that pretreated with second derivative. | 115 |
| 107 | The correlation plots of mannose content between predicted value and actual value of calibration set (a) and validation set (b) | 116 |
| 108 | The residue lines obtained by MWPLSR for mannose content | 117 |
| 109 | The correlation coefficients of mannose from equation 48 that used the combination of wavelength region 1130-1210 nm, 1500-1830 nm, and 2040-2140 nm. | 118 |

LIST OF FIGURES (Cont'd)

| Figure | | Page |
|------------------------|--|-------------|
| 110 | The correlation plots of mannose content between predicted value and actual value of calibration set (a) and validation set (b). | 119 |
| Appendix Figure | | |
| B1 | The gas chromatography calibration curve of standard monosaccharide | 159 |

LIST OF ABBREVIATIONS

| | | |
|--------|---|--|
| Cal | = | calibration set |
| Val | = | validation set |
| MLR | = | multiple linear regression |
| MSC | = | multiplicative scatter correction |
| MWPLSR | = | moving window partial least squares regression |
| NIR | = | near-infrared |
| PLSR | = | partial least square regression |
| R | = | correlation coefficient |
| RPD | = | ratio of the standard error of prediction standard deviation of validation set |
| SD | = | standard deviation |
| SEC | = | standard error of calibration |
| SEP | = | standard error of prediction |

DETERMINATION OF CHEMICAL COMPOSITIONS AND PULP YIELD OF *EUCALYPTUS CAMALDULENSIS* WOOD BY NEAR INFRARED SPECTROSCOPY

INTRODUCTION

Eucalyptus plantation in Thailand is distributed in several parts of the country. Different species are planted depending on area, soil, climate and usage. *Eucalyptus camaldulensis* is an important species in Thailand because it grows very well in various parts of the country, and the wood biomass production is especially excellent for the pulp and paper industries that use *Eucalyptus* wood for raw material. The pulp and paper industries look for wood that has a high pulp yield after the Kraft, or soda process. In addition, for raw wood material the factories need low content in lignin and ash, but high content of pentosan and cellulose. However, the conventional chemical methods are time consuming, require hazardous chemicals, and they are not sufficiently applicable for routine or online pulp quality control in factory. The use of near infrared spectroscopy as an analytical method offers inherent advantages, including: convenient preparation, rapidity, low cost and non-destructive use of the samples, it is an attractive analytical approach for pulp and paper industries to replace conventional chemical method.

The main objective of this study is to investigate the feasibility of predicting pulp yield property and chemical composition of *Eucalyptus camaldulensis* by using NIRs as selection in the tree breeding program of Thailand plantation. The chemical composition divided into three groups by

1. The chemical composition value contained both main components such as holocellulose, α -cellulose, pentosan, lignin, one percentage of Sodium Hydroxide solubility and minor components such as ash, extractive.
2. Monosaccharide value by gas chromatography (GC) method as Glucose, Xylose, Arabinose, Galactose, Mannose.

LITERATURE REVIEWS

1. General Background

The genus *Eucalyptus* contains a remarkably wide range of tree species in regards to adaptation to sites, types of management systems, and variety of uses, both in natural forests and in plantations. *Eucalyptus* can be grown in most of the tropical and temperate climatic regions of the world. (Eldridge, 1997)

In Thailand, prominently *Eucalyptus camaldulensis*, have been established over the last decade to supply woodchips for pulp and paper industries. The *Eucalyptus* wood is a heterogeneous material comprised of cellulose, hemicellulose, lignin, and wood extractive.

2. The Chemical Composition of *Eucalyptus camaldulensis*

The wood cells consist mainly of cellulose, hemicellulose and lignin. A simplified picture is that cellulose forms a skeleton which is surrounded by other substances functioning as matrix (hemicellulose) and encrusting (lignin) materials (Sjostrom, 1993).

2.1 Cellulose

Cellulose is the main constituent of wood. Approximately 40-45% of dry substance in most wood species is cellulose which located predominantly in the secondary cell wall. The length of a native cellulose molecule is at least 5000 nm that corresponding to the chain with about 10,000 glucose units. Cellulose is a homopolymer composed of glucose units which are linked together by β -(1 \rightarrow 4)-glycosidic bonds as show in Figure 1. Cellulose molecules are completely linear and have a strong tendency to form intra and intermolecular hydrogen bonds. Bundles of cellulose molecule are thus aggregated together in the form of microfibrils, in which highly order (crystalline) regions alternate with less ordered (amorphous) regions.

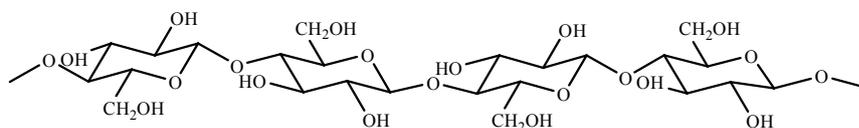


Figure 1 Partial structure of cellulose
Source : Sjostrom (1993)

2.2 Hemicelluloses

Hemicelluloses were originally believed to be the intermediates in the biosynthesis of cellulose. Hemicelluloses are heteropolymer, compose of D-glucose, D-mannose, D-ga lactose, D-xylose, L-arabinose, and small amounts of L-rhamnose. Most hemicellulose has a degree of polymerization of only 200 that are easier hydrolyze by acid than cellulose. The amount of hemicellulose on dry weight of wood is usually

between 20 to 30%. The composition and structure of the hemicellulose in softwoods differ in a characteristic way from those in hardwood. *Eucalyptus camaldulensis* wood is the hard wood.

The major compositions of hemicellulose in *Eucalyptus camaldulensis* wood are glucuronoxylan and glucomannan as show in Figures 2 and 3, respectively. The major component of glucuronoxylan is a D-xylose and glucuronic acid.

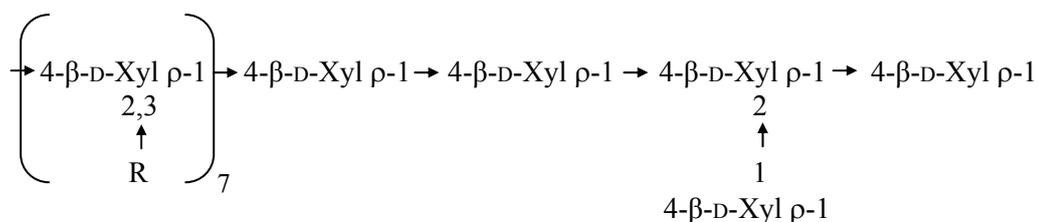


Figure 2 Abbreviated formula of glucuronoxylan.
Source : Sjostrom (1993)

Glucuronoxylan contains a backbone of D-xylose units linked β -(1 \rightarrow 4) with acetyl groups at C-2 or C-3 of the xylose units, average of 7 acetyls per 10 xylose units. The xylan is substituted with side chains of 4-O-methylglucuronic acid units linked to the xylan backbone through a linking (1-2) with an average frequency of approximately an uronic acid group per 10 xylose units.

The major component of glucomannan is D-glucose and D-mannose units link by β - (1 \rightarrow 4) bond.



Figure 3 Abbreviated formula of glucomannan
Source: Sjostrom (1993)

Glucomannan compose of D-glucose and D-mannose linked by β - (1 \rightarrow 4) bond. The glucose: mannose ratio varies between 1:2 and 1:1, depending on the wood species.

2.3 Holocellulose

The carbohydrate portion of the vast majority of plants is composed of cellulose and hemicellulose polymers which are composed of other sugar polymers such as starch and pectin. The combination of cellulose and the hemicellulose are called holocellulose and usually accounts for 65-70 percent of the plant dry weight. These polymers are made up of simple sugars, mainly, D-glucose, D-mannose, D-galactose, L-xylose, L-arabinose, D-glucuronic acid, and less amounts of other sugar such as L-rhamnose and D-fucose. These polymers are rich in hydroxyl groups, which are respectively for moisture sorption through hydrogen bonding.

2.4 Pentosan

Part of the hemicellulose fraction consists of pentose sugars, mainly D-xylose and L-arabinose. The polymers containing these five carbon sugars are referred to as pentosans. Identification of this fraction in plant material has been important to indicate its potential utilization for furan-type chemicals.

2.5 Lignin

Lignin is present in the cellular cell wall, conferring structural support, impermeability. Moreover, it resist against microbial attack and oxidative stress. Structurally, lignin is an amorphous heteropolymer, non-water soluble and optically inactive. It consists of phenylpropane units joined together by different type of linkages. The precursors of lignin are the three aromatic alcohols; coumaryl, coniferyl and sinapyl alcohols as Figure 4, and the structure of lignin as show in Figure 5. Lignins from softwoods are mainly a polymerization product of coniferyl alcohol and are called guaiacyl lignin. Hard wood lignins are mainly syringly-guaiacyl lignin as they are a copolymer of coniferyl and sinapyl alcohol.

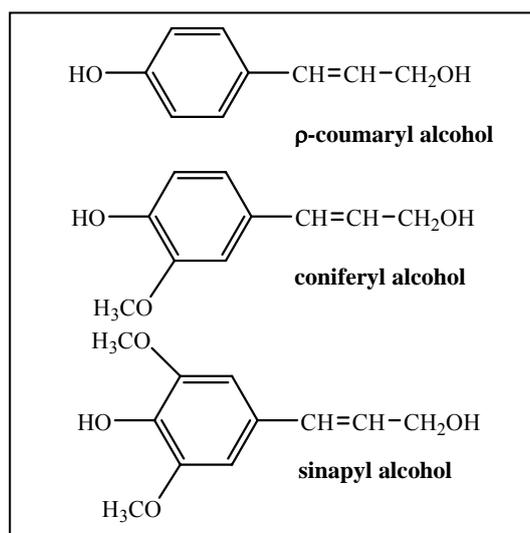


Figure 4 The structure of three precursors of the lignin
Source : Sjostrom (1993)

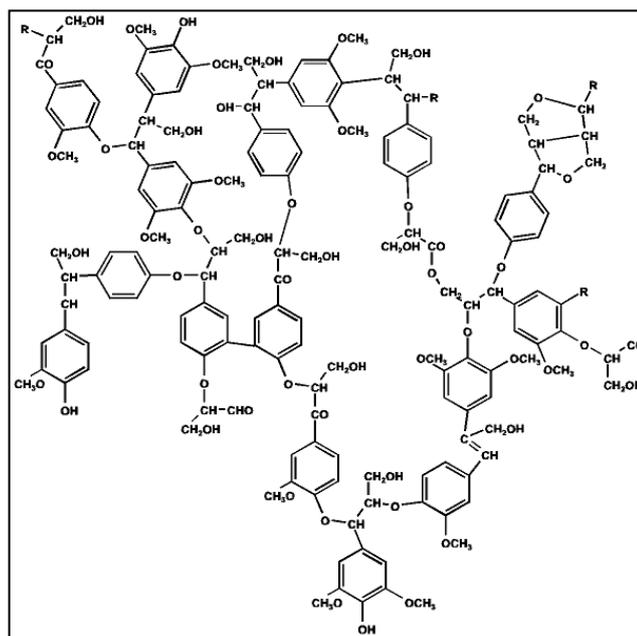


Figure 5 The partial structure of lignin
Source : Sjostrom (1993)

The function of lignin in plants is as an encrusting agent in the cellulose/hemicellulose matrix. It is often referred to as the plant cell wall adhesive. Both lignin and extractive in plant reduce the digestibility in animals. Lignin is also associated with the hemicellulose in some cases forming lignin-carbohydrate complexes that are resistant to hydrolyzed even under pulping condition.

2.6 Wood Extractives

Wood extractives are compounds of diverse nature with low to moderately high molecular weights, which, by definition, are soluble (extract) in organic solvents or water. They are a group of cell wall chemical mainly consisting of fats, fatty acid, fatty alcohols, phenols, terpenes, steroids, resin acids, rosin, waxes, etc. These chemicals exist as monomers, dimmers, and polymers. They impart color, odor, taste, and, occasionally, decay resistance to wood (Biermann, 1996).

2.7 Ash

Ash is inorganic content of a plant, which is an approximate measure of the mineral salts and other inorganic matter in the fiber after combustion at a temperature of $575 \pm 25^\circ\text{C}$. The inorganic content can be quite high in plants containing large amount of silica.

2.8 Monosaccharide

Most of the monosaccharides occur as glycosides and as units in oligosaccharides and polysaccharides and only comparatively few of them are present free in plant. Glucose is the most abundant monosaccharide in nature. It can be

prepared from cellulose and starch by acidic or enzymic hydrolysis of the other aldohexose. Mannose and galactose is important in hemicelluloses. The most common representatives of aldopentoses are xylose and arabinose. Fructose, which represents the only abundant ketose in plant, is presented both free and in a combined state. Fructose is obviously not present in the cell wall polysaccharides of wood.

Analysis of monosaccharides content

Gas chromatography (GC) is a powerful technique for analysis monosaccharide. However, this technique is only suitable for volatile samples. Free monosaccharide cannot inject to gas chromatography directly because of their low volatility. The alditol acetate derivatisation is widely used for determining the composition of monosaccharide mixtures. However, the derivatisation procedure increases the sample preparation time, reagent costs and introduces a degree of error into any quantitative analysis. The alditol acetate derivatisation reaction involves the reduction and acetylating in a series of steps (Figure 6) as follows:

- (a) Reduction of the aldose to an alditol using sodium borohydride.
- (b) Acetylation of the hydroxyl groups on the alditol to form the alditol acetate

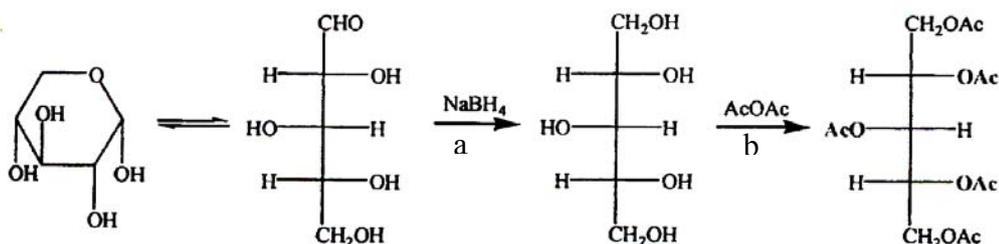


Figure 6 The alditol acetate derivatisation reaction of xylose
Source : Chaikumpollert (2003)

3. Pulping

Pulping is the process of converting wood for papermaking. Processes range from purely mechanical pulping, in which the wood is ground into fibers by disk refiners or grindstones, to chemical pulping, in which the fibers are separated by chemical degrading and dissolving the lignin that binds them together in the tree. Thus, chemical pulping methods rely on the effect of chemicals to separate fibers, whereas mechanical pulping methods rely completely on physical action. Moreover those chemicals are involved, the lower the yield and lignin content since chemical action degrades and soluble components of the wood, especially lignin and hemicellulose. On the other hand, chemical pulping yields individual fibers are not cut and provide the strong paper since the lignin, which interferes with hydrogen bonding of fibers is largely removed (Biermann, 1996).

The objective of chemical pulping is to remove lignin not only from fiber walls but also from the middle lamella so that the wood fibers can separate. Chemical pulping accounts 70% of the total worldwide production, and currently about 80% of chemical pulps is produced by the dominant kraft process. Present, soda pulping is used in minor process. However, soda pulping has not odorous effect. Kraft process can be replaced by soda anthraquinone processes which process a marked capability of accelerating the delignification which at the same time stabilizing the polysaccharides toward alkaline degradation. The choice of pulping process is determined by the desired product, wood species available, and economic considerations.

3.1 Reaction of Lignin

The general nature of Kraft and Soda pulping reaction with lignin is nucleophilic with respect to delignification, an essential aspect is the different behavior and stability of the various types of linkages between phenyl propane units in lignin. Depolymerization of lignin typically depends on the cleavage of all types of aryl ether linkages. During delignification, some condensation of lignin units also occurs, leading to fragments with increased molecular mass and reduced solubility. The most important reductions are as follows:

- Cleavage of β -aryl ether linkages in nonphenolic structures
- Cleavage of β -aryl ether linkages in free phenolic structures
- Cleavage of α -aryl ether bonds in free phenolic structures
- Demethylation reactions
- Condensation reactions

The mechanism of reaction can be explained as follow:

3.1.1 Cleavage of β -aryl ether linkages in nonphenolic structures

In etherified *p*-phenolic structures the β -aryl ether linkage is cleavage by hydroxide ions according to mechanism shown in Figure 7. The reaction proceeds via an oxirane intermediate which is subsequently opened with formation of an α , β -glycol structure. This reaction promotes efficient delignification by the fragmenting the lignin and by generating new free phenolic hydroxyl groups.

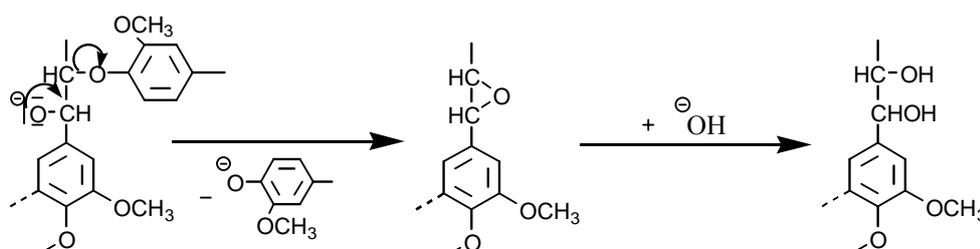


Figure 7 Cleavage of β -Aryl ether bonds in nonphenolic phenylpropane units during Soda pulping.

Source : Sjostrom (1993)

3.1.2 Cleavage of β -aryl ether linkages in free phenolic structures

In the first step of the reaction involves the formation of a quinone methide from the phenolate anion by the elimination of a hydroxide, alkoxide, or phenoxide ion from the α -carbon (Figure 8). The subsequent course of reaction depends on whether hydrogen sulfide ions are present or not. In the latter case (soda pulping), the dominant reaction is the elimination of formaldehyde and a styryl aryl ether structure without β -ether bonds (Figure 9).

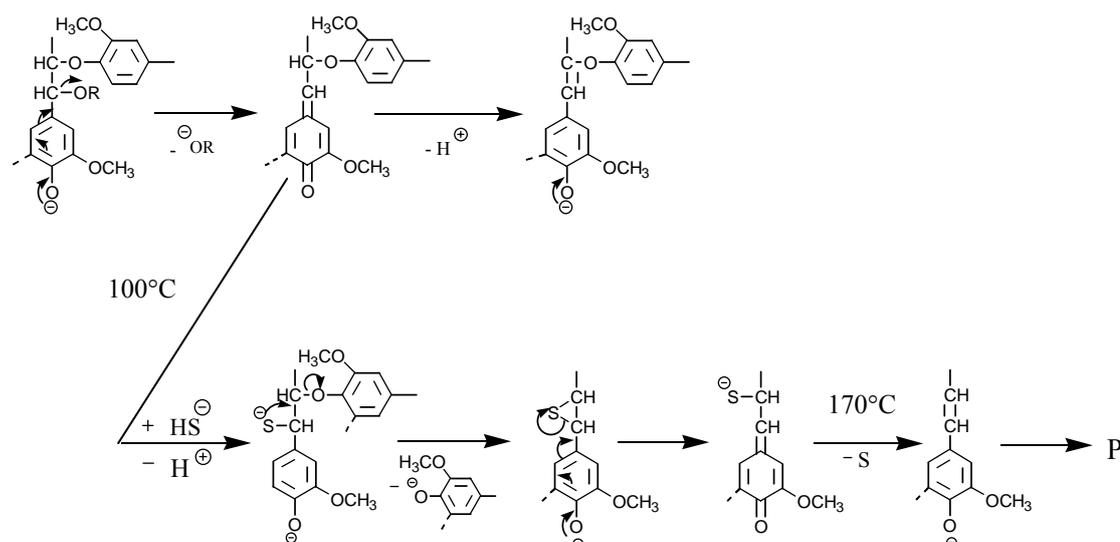


Figure 8 Main reactions of phenolic β -aryl ether structures during alkali(soda) and Kraft pulping. R=H, alkyl, or aryl group.

Source : Sjostrom (1993)

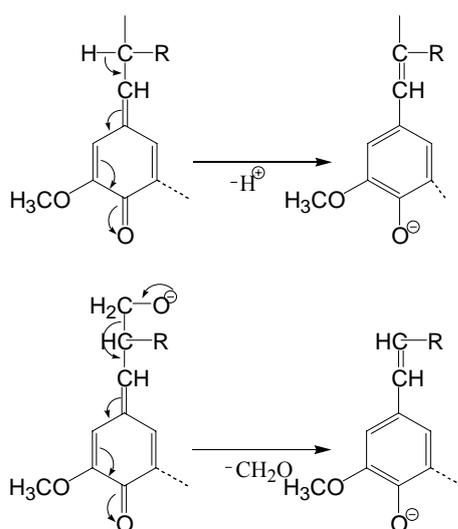


Figure 9 Elimination of proton and formaldehyde from the quinone methide in intermediate during alkali pulping.

Source : Sjostrom (1993)

3.1.3 Cleavage of α -aryl ether linkages in free phenolic structures

The α -Ether bonds in phenolic phenylcoumaran (Figure 10) and pinoresinol structures are readily cleaved by hydroxide ions, usually followed by the release of formaldehyde. Only in the case of open α -aryl ether structure does this reaction result in the fragmentation of lignin. In contrast, the α -ether bonds are stable in all etherified structures.

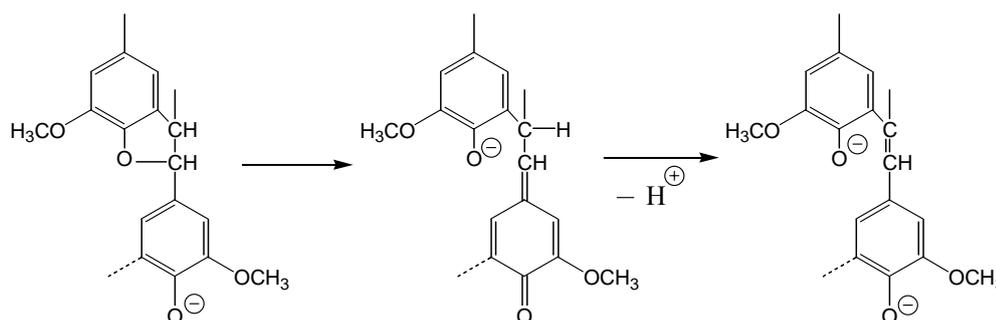


Figure 10 Cleavage of α -aryl ether bonds in free phenolic structures
Source : Sjostrom (1993)

3.1.4 Demethylation reactions

Lignin is partially demethylated by the action of hydrogen sulfide ions forming methyl mercaptan which is convertible to dimethyl disulfide by reaction with another methoxyl group. In the presence of oxygen, methyl mercaptan can be oxidized further to dimethyl disulfide (Figure 11), because the hydroxide ions are less strong nucleophiles than hydrogen sulfide ion. Methyl mercaptan and dimethyl sulfide are highly volatile and extremely malodorous, causing an air pollution problem.

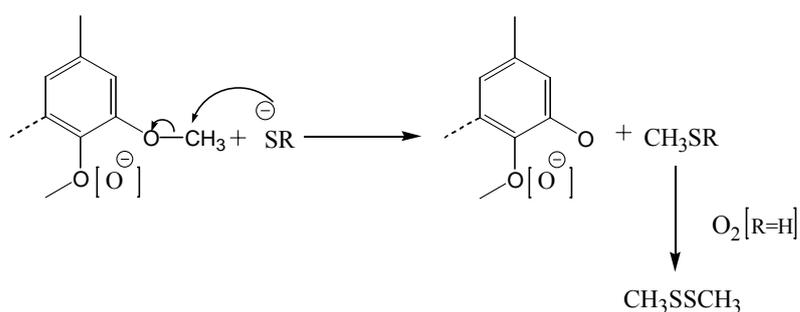


Figure 11 Cleavage of methyl aryl ether bonds with simultaneous formation of methyl mercaptan (CH_3SH), dimethyl sulfide (CH_3SCH_3), and dimethyl disulfide (CH_3SSCH_3) during kraft pulping. $\text{R} = \text{H}$ or methyl group.
Source : Sjostrom (1993)

3.1.5 Condensation reactions

A variety of condensation reactions are known to occur in pulping. In these reactions a phenolate unit is added to the quinone methide intermediate, forming a new carbon-carbon bond by an irreversible release of a proton. The major case occurs at the C-5 position of phenolic units, forming an α -5 linkage (A). The second case (B) illustrates similar condensation between the 1- and α -carbons with simultaneous removal of the propane side chain. The third reaction (C) involves formaldehyde released from the γ -carbinol groups (Figure 12) and also leads ultimately to diarylmethane structure.

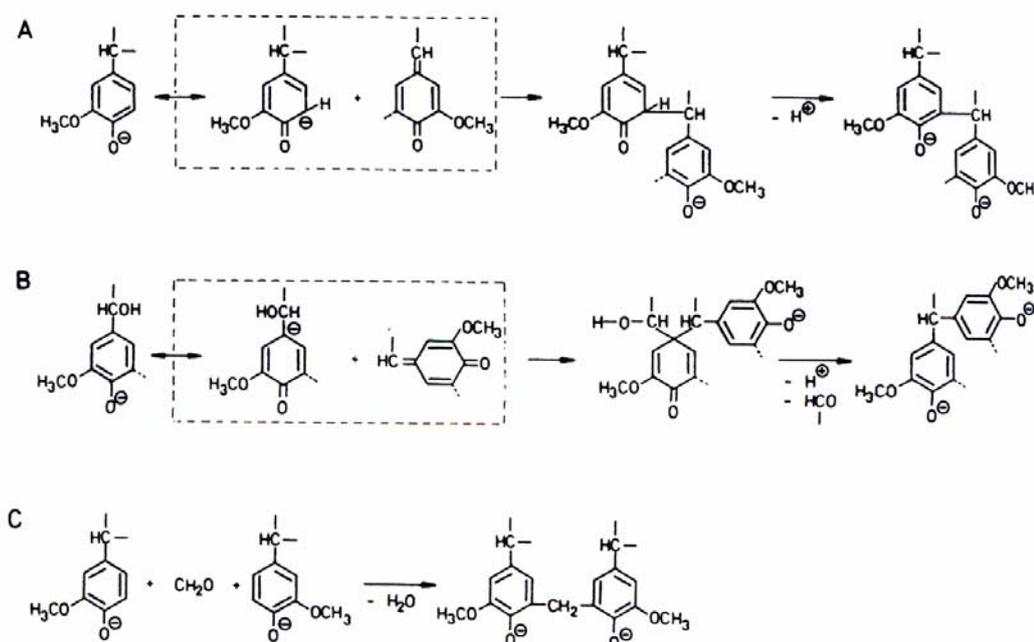


Figure 12 Examples of condensations reactions during alkali and kraft pulping.
Source : Sjostrom (1993)

3.2 Reaction of the Polysaccharides

The most important alkali-catalyzed reactions for the loss of polysaccharides and the reduction of their chain length in soda pulping are the peeling reaction and the stopping reaction. The mechanism and explanation of both reactions show in Figure 13.

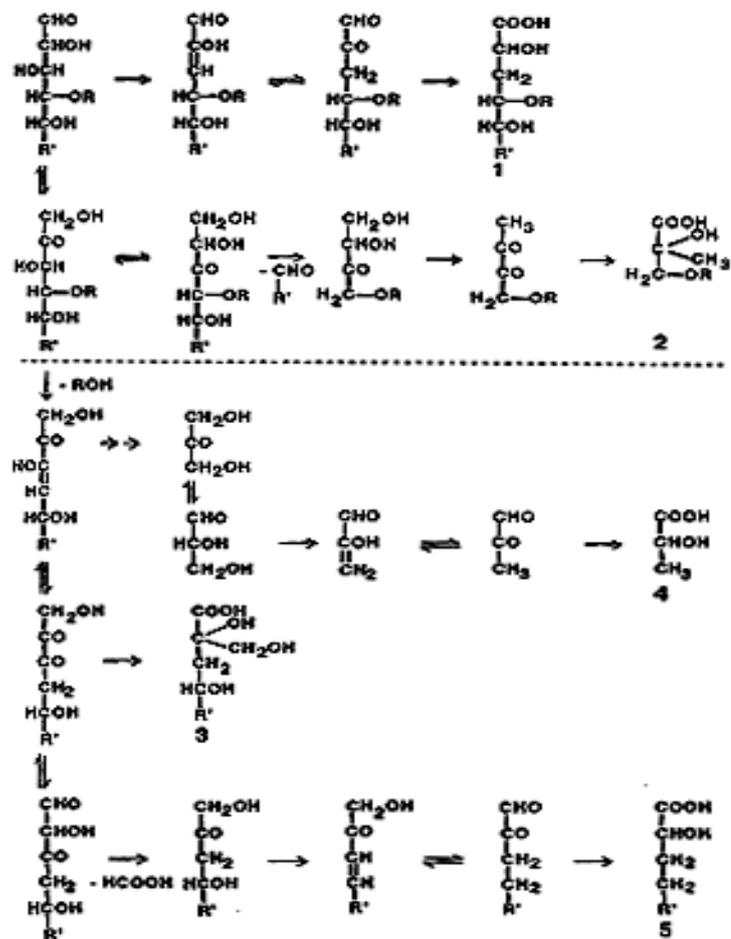


Figure 13 Peeling and stopping reactions of polysaccharides.
Source : Sjostrom (1993)

4. Descriptions of Chemical Method in Pulp and Paper Industry

Total pulp yield is indicated the percentage of all material that will not dissolve in the pulping process.

Screen pulp yield is the cellulose fiber passed the screening process. This fiber is pulp paper for papermaking.

Holocellulose is the term applied to the total carbohydrate fraction in extractive-free wood. It is isolated from these materials by cautious delignification, as with acidified chlorite. The separation will always be somewhat in error. Because as the lignin content becomes very small, so some carbohydrates will be attacked and lost.

α -Cellulose is the portion of cellulosic material which resists in sodium hydroxide solution. α -Cellulose consists chiefly of cellulose, therefore the content was related to the percentage of pulp yield.

The determination of pentosan is based upon the production of furfural from the xylose and arabinose units in hemicellulose.

Lignin is usually determined as the residue left on hydrolysis with strong acid. In the alkaline process, lignin is removed during chemical pulping operation. It appears in the spent liquors as alkali lignin. In the acid or neutral sulfite process, the lignin is sulfonated to ligninsulfonic acids, and occurs in the spent liquors as calcium, ammonium, magnesium, or sodium salts of these acids. By-product lignin has found important uses as dispersing agents, oil well drilling mud additives, and as a commercial source of vanillin.

Ash content is the percentage of inorganic residue of completed combustion of cellulosic material. It is an indication of the amount of mineral salts and inorganic foreign matter in wood and pulp. It is also indication of the filler, coating, pigmentation, and chemical additive content in a paper sheet.

Extractives are not part of the cell wall structural elements. It are the basis for some by-products of the cooking process.

One percent of NaOH solubility values are indicated the low molecular of cellulose that contained in wood material. It can refer to yield value in pulping process.

5. Principles of Near Infrared Spectroscopy

Vibration spectroscopy is based on the concept that atom-to-atom bonds within molecules vibration with frequencies that may be described by the laws of physics. When these molecular vibrators absorb the light of a particular frequency, they are excited to higher energy level. The lowest, or fundamental, frequencies of any two atoms connected by chemical bonds may be roughly calculated by assuming that the bond energies are arise from the vibration of a diatom harmonic oscillator, and it obeys Hooke's Law.

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

where ν = the vibration frequency
 k = the classical force constant
 μ = the reduce mass of the two atoms

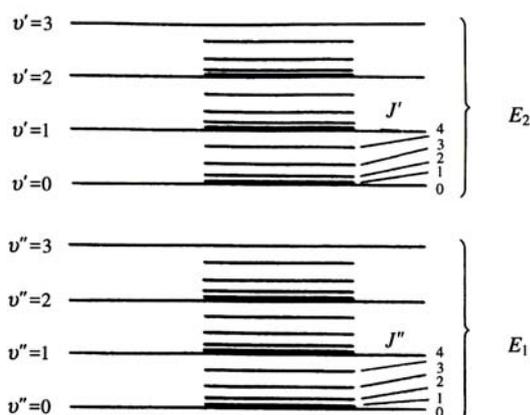


Figure 14 Energy levels of a molecule. E_1, E_2 : electronic energy levels. $v''=0,1,2,\dots,$
 $v'=0,1,2,\dots$: vibration energy levels. $J''=0,1,2,\dots, J'=0,1,2,\dots$ rotational
 energy levels.

Source : Ozaki (2003)

Figure 14 shows the energy levels of a molecule. Since intervals between energy levels in the electronic state are quite large, a considerable amount of energy is required to cause transition between electronic states. Visible or ultraviolet light has energy that corresponds to this required amount of energy. The mid-IR region relates essentially to vibration states of molecules. The energy that is necessary to cause vibration transitions is much small, approximately one-tenth of the energy that is needed to cause an electronic transition. The infrared region may be divided in three regions: near, middle and far infrared presented in table 1

Table 1 Divisions of the infrared region (Osborne, 1986)

| Region | Characteristic transition | Wavelength range (nm) | Wave number range (cm^{-1}) |
|-----------------------|----------------------------|-----------------------------|--|
| Near infrared (NIR) | Overtones and combinations | 800-2500 | 14300-4000 |
| Middle infrared (MIR) | Fundamental vibrations | 2500- 5×10^4 | 4000-200 |
| Far infrared (FIR) | Rotations | 5×10^4 - 10^6 | 200-10 |

The ‘near-infrared’ region (NIR) mean ‘the closest’ region of the infrared region to the visible light region. The NIR region covers the interval between approximately 800 and 2500 nm as present in Figure 15.

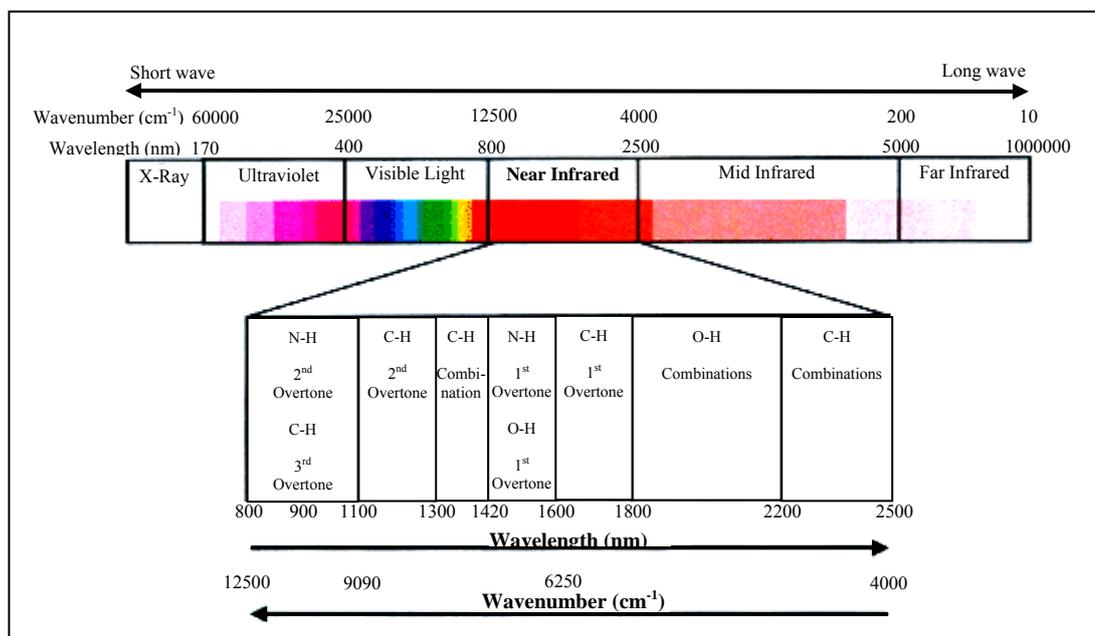


Figure 15 Present various region of spectroscopy.

Source : Anonymous(2001)

The NIR region is not associated with any fundamental transitions, such as transitions between electronic states for the visible region or transitions between vibration states. That band is arising from overtones and combination modes of molecular vibration.

Overtone and combination modes are so-called “forbidden transitions”, which are transitions that should not be observed in harmonic oscillator approximation. In reality, the fact that the NIR region is a forbidden transition is what makes this region unique and markedly different from the other regions. The reason why the NIR region is precious is only the NIR region serves as a window to light, in one sense, while neighboring regions (the violet region, the visible region, the IR region and the far IR region) are all regions where a number of allowed transitions appear. In general, the NIR region transmits light well. In the other words, absorption is weak in the NIR region. However, the extent of the weakness of absorption is different depending on the range of the NIR region.

The NIR region may be divided into three regions: region I (800-1200 nm; 12500-8500 cm^{-1}), region II (1200-1800 nm; 8500-5500 cm^{-1}) and region III (1800-2500 nm; 5500-4000 cm^{-1}). Region I is sometimes called ‘the short-wave near infrared (SWNIR) region’ or ‘the Herschel region.’ This region is the region in which second and third overtones and combination modes of XH (X=O, N, C) stretching vibrations are observed, and which exhibits a very high transparency. In some cases,

bands due to electron transitions appear in this region. In region II, that observes first overtones of XH stretching vibrations and various types of combination modes of XH stretching vibrations. Region III is a combination mode region. Overtone appearing, if any in this region, are at most second overtones of C=O stretching vibrations. The 'permeability' of region III is comparatively poorer (Chalmers, 2000).

Quantitative analysis of NIR spectroscopic data is based on Beer-Lambert law, which states that a linear relationship exists between the molar concentration of a substance and the absorbance.

$$A = \log \frac{1}{T} = \log \frac{P_0}{P} = \epsilon bc \quad (2)$$

Where A is absorbance, P_0 is the power of the incident, P is the power of the transmitted radiation, ϵ is the molar absorptivity, b is the path length of the sample and c is the concentration of the compound in solution. (Siesler *et al.*, 2002)

Modern spectroscopic practice has progressed far beyond the simple use of Beer's law to relate the absorbance of an analyte at a characteristic wavelength to the concentration of that analyte. To allow accurate analysis in the other form of noise, inhomogeneous samples, and other phenomena that can affect spectroscopic reading, a variety of sophisticated mathematical techniques (i.e. chemometrics) have been developed that attempt to extract the analytical information from the spectroscopic data.

5.1 Concept of NIR Analysis

Typical steps in develop a near infrared analysis procedure include optimize the instrument and sampling system. The primary analysis method, collect spectra for the calibration set, process the spectra data to improve resolution and remove sampling artifacts, apply a modeling technique to the data which is usually a linear model such as multiplicative linear regression (MLR), Partial least square regression (PLSR), and validated technique by predicting analytical values for sample independent of the calibration set. Statistic test may be applied to detect outlier both in the calibration and validation sets. This diagram illustrates the analysis.

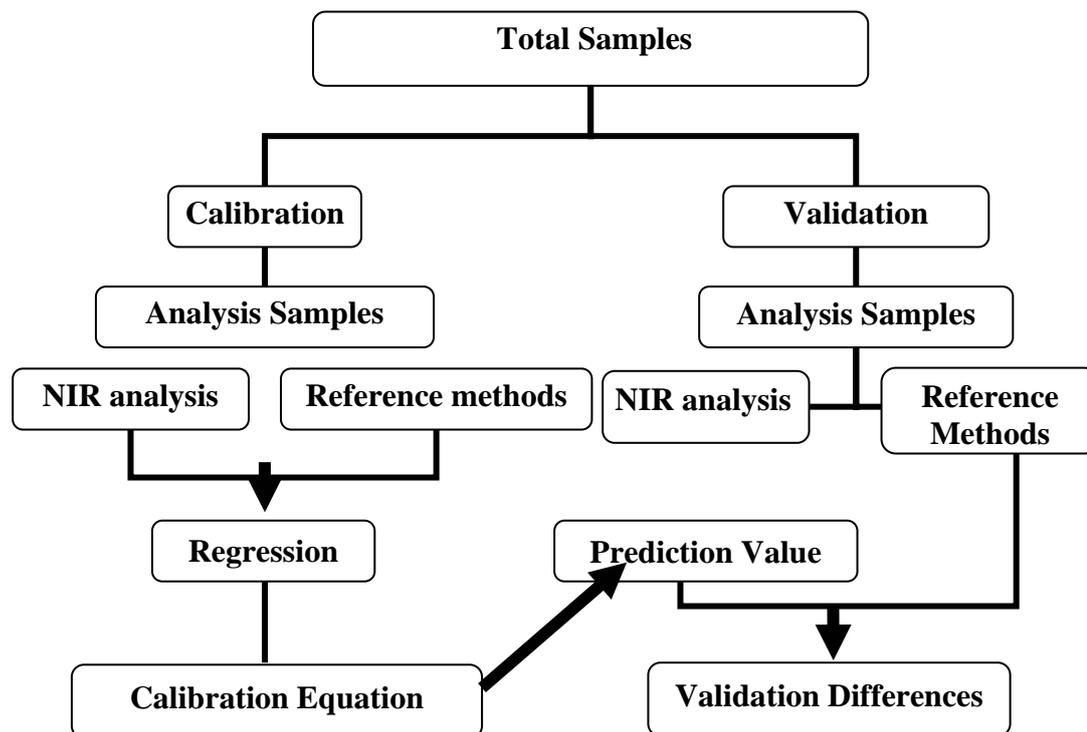


Figure 16 Diagram of NIR analysis concept.
Source : Suwonsichon (2002)

5.2 Instrument

The spectrometer includes a light source, a means of dispersing the light in the near infrared region, optics for focusing the incident light to reflect or transmitted light from the sample onto a detector.

Several types of instrumentation are in use, including filter-based instruments (photometers), scanning-grating-based instruments (dispersive spectrophotometers), Fourier transform spectrophotometers, diodes, and acousto-optically tuned filter (AOTF) based instruments. Figure 17 shows the schematic diagram of NIR spectrometer using grating type.

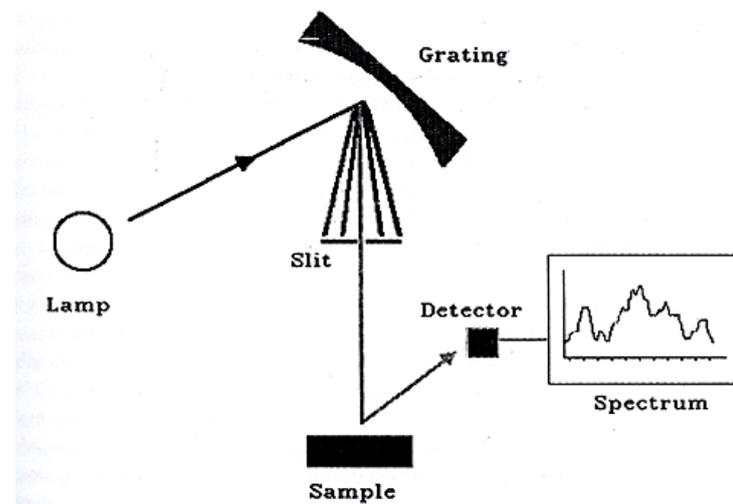


Figure 17 NIR spectrophotometer: Schematic diagram of a dispersive scanning grating near-infrared spectrophotometer.

Source : Connors (1995)

The light source is a tungsten halogen current regulated lamp enclosed in a reflective sphere to enhance infrared emission of the filament image. A refractive lens system forms the image at a variable entrance slit that is controlled by a stepper motor. The variable entrance slit ensures constant light output from the monochromatic grating throughout the entire scan range. Light leaving the entrance slit strikes a mirror that reflects the multichromatic light beam onto a concave grating mirror that is stepper motor driven. The uniformly illuminated grating, which is rotated in increments on an axis about its vertex, emits monochromatic light as a function of wavelength across the scan range.

The NIR instrument begins with: the beam of light entering the optical head passes through a stepper motor driven Filter Wheel containing three blocking filters. The system program commands the required blocking filter to be presented as a function of the wavelength being scanned. A red glass filter is used from 700 to 1100 nm, a silicon filter from 1100 nm to 1940 nm, and a germanium filter from 1941 to 2500 nm. Light exiting the filter wheel is beamed onto a continuously rotating chopper wheel. The reference beam that passes through the chopper wheel is reflected by the reference mirror and directed toward the side of the integrated sphere producing only specular light within the sphere. This reflected light sensed by the detectors in the sphere is the "Reference" signal. When a mirror segment on the wheel is placed in the path of the light beam, the beam is directed onto the sample. The 'sample' reading is the result of diffuse reflectance light, which is affected by the amount of absorption in the sample surface. The light path is shown in Figure 18 and the instrument layout of NIR is shown in Figure 19.

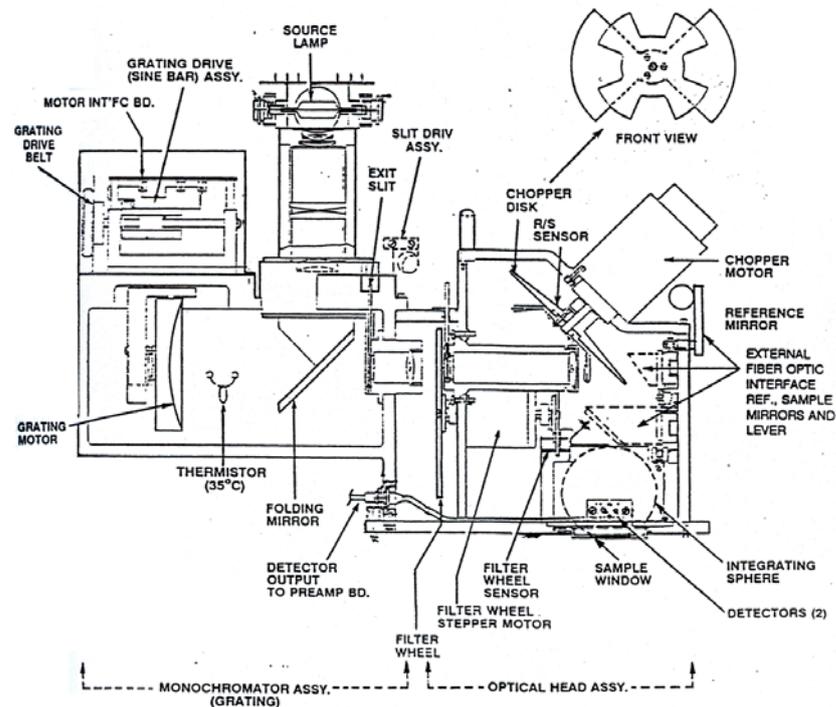


Figure 18 Optical Head and Monochromator

Source : Anonymous (2001)



Figure 19 Near Infrared Spectroscopic Instrument of InfraAlyzer 500 (Bran+Luebbe)

6. Sample Presentation

“Sample presentation” or how to set sample to a NIR Instrument is one of the important factors affecting NIR Instruments. Figure 20 illustrates some sample presentations known as “transmission”, “reflection”, “tranflection” and “interaction”. In the case of reflection, incident light illuminates the surface of the sample and the diffusely reflected light from the surface may be detected. The sample should be opaque, a powder sample, and have more than 1 cm depth (Siesler *et al.*, 2002).

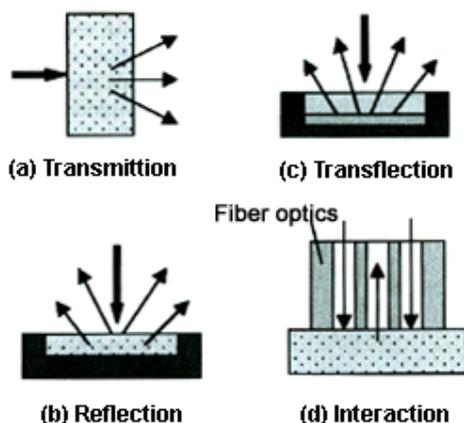


Figure 20 Sample presentations of transmission(a), reflection(b), tranfection(c) and interaction(d)

Source : Siesler (2002)

7. Chemometric Methods for Quantitative Analysis

NIR information can not be used to determine analyze concentrations directly because of the way in which near infrared radiation passes into, and reflect from the sample. Chemometric are necessary to make the calibration model.

7.1 Modern Pretreatments Techniques of NIR Spectra

NIR bands are broader than in mid-infrared and spectra are considerably more complex. Due to this complexity, NIR spectroscopy uses of the data pretreatment.

7.1.1 Multiplicative Scatter Correction (MSC)

Multiplicative Scatter Correction (MSC) was developed to reduce the effect of scattered light on diffuse reflection and transmission of NIR spectra. Scattering should have a multiplicative effect on reflection (and transmittances) spectra. That is, the observed spectra will contain broad band, changing back ground from differential scattering at each wavelength.

This method assumes that the wavelength dependency of the light scattering is different from that of the constituent absorption. Theoretically, by using data from many wavelengths in the spectrum, it should be possible to separate the two wavelengths. MSC is a powerful preprocessing method for removing additive and multiplicative differences in a spectrum. In Figure 21A, it is apparent that the largest source of variance within the 40 samples is derived from scattering. Figure 21B presents the plots of the spectra are pretreated with MSC (Burns, 2001).

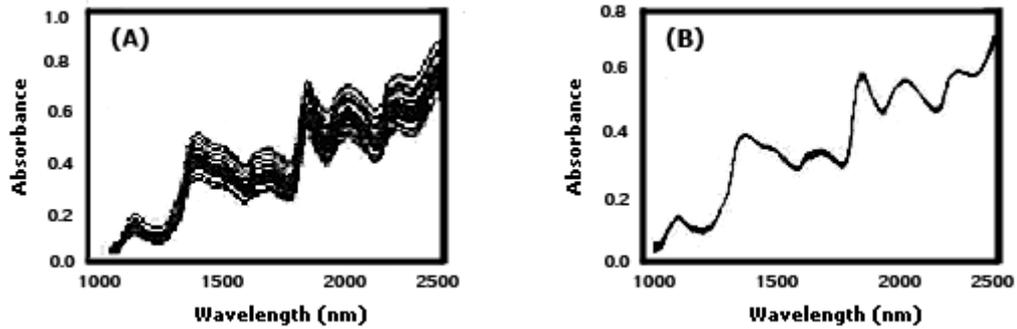


Figure 21 Forty corn calibration samples: (A) untreated and (B) MSC applied to 40 calibration samples.

Source : Boysworth and Booksh (2001)

The idea of MSC lies in the fact that light scattering has the wavelength dependence different from that of chemically based light absorption and light scattering. MSC corrects spectra according to a simple linear univariate fit to a standard spectrum and estimated by least square regression using the standard spectrum. As a standard spectrum, a spectrum of a particular sample or an average spectrum is used, where the scattering at the i th wavelength of a sample can be modeled by

$$x_i = a + b\bar{x}_i \quad (2)$$

With a and b being constant for all i wavelengths in sample. The scatter-corrected data is determined by the scale deviations about the regression. x_i is absorbance and \bar{x}_i is an average spectrum for all i wavelengths in sample (Burns, 2001).

$$x_{i(MSC)} = \frac{(x_i - a)}{b} \quad (3)$$

7.1.2 Derivative Methods

Derivative methods have long been used in NIR spectroscopy as pretreatment methods for resolution enhancement, as well as baseline correction. The second derivative is the most often used one. The superimposed peaks in an original spectrum turn out as clearly separated downward peaks in the second derivative spectrum. Another important property of the second derivative method is the removal of the additive and multiplicative baseline variations in an original spectrum. On the other hand, a drawback in the derivative methods is that the spectra to noise ratio deteriorates every time a spectrum different. A simple formula for calculating second derivative is give by:

$$d^2 A_i = d(A_{i+k} - A_{i-k}) = A_{i+2k} - 2A_i + A_{i-2k} \quad (4)$$

It is very difficult to discuss the band shift in the NIR region of the original spectra, but the calculation of the second derivative makes number band clearly detectible. The typical second derivative spectra presented in the Figure 22.

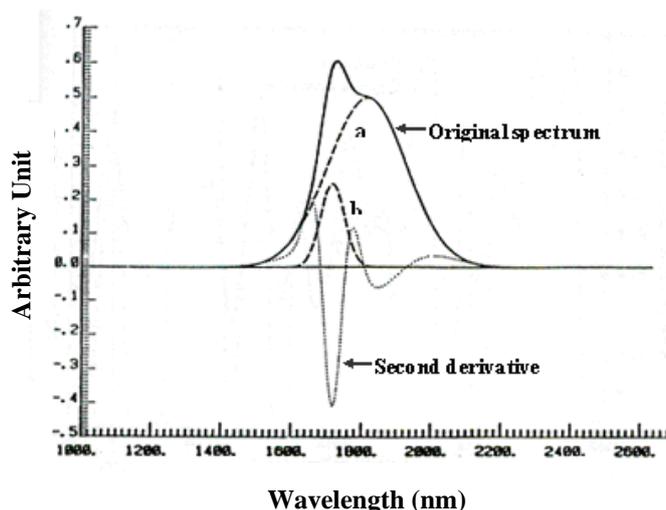


Figure 22 Original Spectrum that combine from two Gaussian curves (a and b) compared with the Second derivative spectrum
Source : Hruschka (1990)

7.2 Multivariate Calibration Methods

7.2.1 Multiple Linear Regressions (MLR) Technique

The purpose of multiple regressions is to find the relationship between several independent, or predictor variables, and a dependent, or criterion variable. Multiple linear regressions (MLR) is a quantitative calibration model that allows one to establish a relation between NIR spectra and quantifiable sample properties, e.g. contents of components such as lignin content, holocellulose, etc. (Anonymous, 1998)

The basis for an MLR calibration are two sets of data: the set of reflectance values measured at the different wavelengths for the calibration set samples, and of the reference property values for the corresponding samples. If a transformation is selected for the calibration, e.g. absorbance, or second derivative, the transformed values will be used instead of the measured ones. The basic procedure of a linear regression consists of plotting the measured, or transformed, values at a given wavelength against the reference property values for all spectra in the calibration set. Then least square fitting is performed to determine the regression line. This regression line is the line that minimizes the sum of the squares of the vertical distances of the property values from the line. The regression line is calculated by using the following equation:

$$y = F_0 + Fx \quad (5)$$

y = property value

F_0 = interception of the regression line (the value of y when x is zero)

F = slope of the regression line (called regression coefficient)

x = measured, or transformed value at a specific wavelength.

This equation, using one wavelength to calculate the property value, would be enough in the ideal case where Lambert-Beer's law applies, e.g. in case of UV/VIS spectra of diluted solutions. For NIR analysis of solid or liquid samples, however, there will always be non-linearity, for example because of interactions between constituents or particle size effects. For this reason, a multiple linear regression will be required, using several wavelengths and regression coefficient. The corresponding equation can be written:

$$y = F_0 + F_1x_1(\lambda_1) + F_2x_2(\lambda_2) + F_3x_3(\lambda_3) + F_4x_4(\lambda_4) + \dots \quad (6)$$

For making the optimum calibration, it must consider statistic result parameter, e.g. multiple correlations coefficient (R), standard error of estimate (SEC), standard error of prediction (SEP) etc.

The multiple correlations coefficients R are a measure of the agreement between the reference property values and the predicted values. Or, in other words, it is a measure of the correlation between the absorbance, or transmittance/reflectance, values at the chosen wavelengths and reference values. The equation for multiple correlations coefficients is calculated as equation 7.

$$R = \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2 \sum_{i=1}^n (x_i - \bar{x})^2}} \quad (7)$$

The maximum value is 1, which indicates a perfect correlation. The Standard error of calibration (SEC) indicates how well the calibration equation fits the data. It is indicated the error that we can expect if we use this calibration model for predictions. SEC calculated from the difference between the reference properties values (wet chemical values) and the predicted property values calculated for the calibration set spectra as equation 8 (Schimleck *et.al.*, 2003).

$$SEC = \sqrt{\frac{\sum_{i=1}^{m_p} (\hat{y}_i - y_i)^2}{m_c - k - 1}} \quad (8)$$

A good sample set with an appropriate distribution of property values will often have an SEC is 0.1 times the Standard Deviation of the reference property values (Anonymous, 1998).

7.2.2 Partial Least Square Regression (PLSR) Technique

There are many methods to select the range of wavelength before making calibrations by PLSR. In this research, the whole wavelength is exported to moving window partial least square regression by in-house-written programs in MATLAB (ver.5.3: The Math Works, U.S.A.) for wavelength selection (Kasemsumran, 2005).

7.2.2.1 Moving Window Partial Least Squares Regression (MWPLSR)

MWPLSR was proposed in Ozaki group of Kwansai Gakuin University as a new wavelength interval selection method for multicomponent spectra analysis. It has been demonstrated theoretically that MWPLSR can provide an efficient calibration model with high stability against the interferences from non-composition related factors. In MWPLSR, a spectral window X_i (matrix $m \times h$) is constructed and moved over the whole spectral region (matrix $m \times n$). The position of this window is used for indicating the starting position of the window. At each window ($i = 1, 2, \dots, (n - \text{window size } h + 1)$), spectral data (X_i) are collected and employed for developing PLS models with different numbers of PLS factor (latent variables, LVs). Then, the log of sums of squared residues ($\log(\text{SSR})$) are calculated with these PLS models and plotted as a function of the position of the window. A single line plot of $\log(\text{SSR})$ values corresponds to the values of $\log(\text{SSR})$ at a certain model dimensionality in the corresponding window position. In this residual lines plot, the informative spectral regions for an analyte can be selected from the spectral regions where the residual lines show low values of $\log(\text{SSR})$. A scheme of MWPLSR calculation processes is given in Figure 23. In this scheme, y is the concentrations of the analyte and b_i is the regression coefficient estimated using PLS with a number of PLS factor.

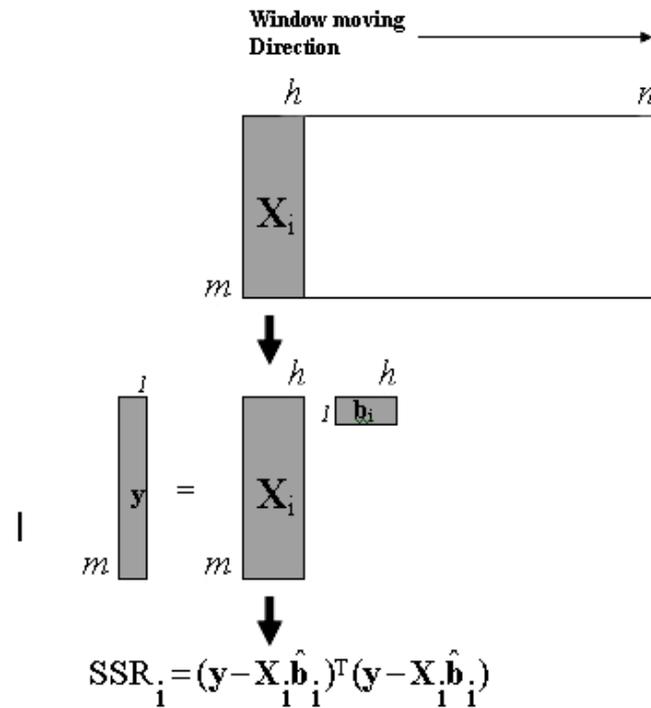


Figure 23 Scheme for explanation of MWPLSR

Source : Kasemsumran(2005)

7.2.2.2 Partial Least Square Regression (PLSR) Technique

PLS regression is a multivariate data analytical technique that lies in the construction of factors, as linear combinations of the original spectral data (x -values), and employs only these factors in the regression equation. PLS constructs the factors that capture as much of the variations in the spectral data as possible, and it takes into account whether, they correlate with the reference data or not. The aim of PLS regression is to reduce the quantity of spectral data, but only the most relevant part of the x -variation is used in a regression for predicting y . A multivariate calibration model can be described as follows:

$$y = Xb + e \quad (9)$$

Where X is an $m \times n$ matrix, m and n are the numbers of samples and variables, respectively. Y is the response vector with the dimension of $m \times 1$, b is the regression coefficient vector, and e is the error vector. When the linear model (1) is considered, the matrix X can be decomposed as follows:

$$X = t_1 p_1^T + t_2 p_2^T + \dots + t_k p_k^T + F \quad (10)$$

Where t_i , p_i are the scores and loadings and F is the residual matrix. Here, the superscript T denotes transpose. Then,

$$X = \sum_{i=1}^k t_i p_i^T = TP^T \quad (11)$$

Where $T = [t_1, t_2, \dots, t_k]$ and $P = [p_1, p_2, \dots, p_k]$. Each score t_i is a combination of the column vector of matrix X and W is the appropriate weight matrix, that is

$$T = XW \quad (12)$$

The next step is to find the regression coefficient q by regressing y onto T . The regression model can be written as

$$y = Tq^T + e \quad (13)$$

The model (9) can also be rewritten as:

$$y = TP^T b + e \quad (14)$$

Thus, the regression coefficient b used in the linear PLS predictor can be calculated using the equation

$$b = W(P^T W)^{-1} q^T \quad (15)$$

The optimal component number is decided for containing in a model, and it is denoted by j (normally, $k > j$). The metrics T , P and W can be expressed as; $T = [t_1, t_2, \dots, t_j]$, $P = [p_1, p_2, \dots, p_j]$, and $W = [w_1, w_2, \dots, w_j]$. Finally, the equation below can be reached;

$$y = T_j P_j^T b + e \quad (16)$$

Generally, PLS regression has the objective to apply for one single y -variable so called PLS1 regression. It can be also modified for analyzing two or more y -variables simultaneously, namely PLS2 regression. However, PLS2 is easily affected by non-linearity in the X - y relationship and random noise. Therefore, the use of separate PLS1 modeling for each y -variable may be the best way to obtain optimal calibration models.

Two further quantities which are usually calculated as measures of the goodness of the fit line are the residual standard deviation and the correlation coefficient.

$$\sigma = \sqrt{\sum_{i=1}^n (y_i - a - bx_i)^2 / (n - 2)} \quad (17)$$

The correlation coefficient measures the extent to which the fitted straight line relationship explains the variability in the y -values. In fact, R which lies between 0 and 1, is the proportion of the total variance in the y -values explained by the fitted line.

$$R = \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2 \sum_{i=1}^n (x_i - \bar{x})^2}} \quad (18)$$

X_{ji} is the observed value of the j th X variable (the absorbance value at wavelength channels j of sample i), when $i = 1, 2, \dots, n$ and $j = 1, 2, \dots, m$.

$$\begin{array}{c}
 \begin{array}{|c|} \hline n \\ \hline \mathbf{X} \\ \hline m \end{array}
 =
 \begin{array}{|c|} \hline 1 \quad k \\ \hline \mathbf{T} \\ \hline m \end{array}
 \begin{array}{|c|} \hline 1 \\ \hline \mathbf{W}^T \\ \hline k \end{array}
 +
 \begin{array}{|c|} \hline n \\ \hline \mathbf{F} \\ \hline m \end{array}
 \\
 \\
 \begin{array}{|c|} \hline \mathbf{y} \\ \hline m \end{array}
 =
 \begin{array}{|c|} \hline 1 \quad k \\ \hline \mathbf{T} \\ \hline m \end{array}
 \begin{array}{|c|} \hline \mathbf{q} \\ \hline 1 \end{array}
 +
 \begin{array}{|c|} \hline \mathbf{e} \\ \hline m \end{array}
 \end{array}$$

Figure 24 The concept of PLS calculation.
Source : Kasemsumran (2005)

7.3 Validation Methods

In order to select the optimal number of PLS factors and to determine the prediction ability of a PLS calibration model, the various validation methods have been proposed. All validation methods used in this study are described below.

7.3.1 Test Set Validation

Test set validation is based on testing the calibration model by splitting the sample data into two sets; one is the calibration set for developing a model and the other is the test set for testing the model. Then, the standard error of prediction (SEP) can be calculated, that is (Schimleck, 2003)

$$SEP = \sqrt{\frac{\sum_{i=1}^{m_p} (\hat{y}_i - y_i)^2}{m_p - 1}} \quad (19)$$

where \hat{y}_i and y_i are the predicted and reference values for the test set, respectively. The number of samples in the test set is defined as m_p .

By using only one test set for validating the model, the samples contained in the test set should cover the relevant range of samples. Note that the calibration model will show the different results when the different test sets are used for validation. (Kasemsumran, 2005)

A useful method for evaluation of a calibration involves the RPD. That is the ratio of standard error of prediction (SEP) to standard deviation (SD) of validation sample set. If SEP values are similar to the SD value, it means that the instrument is not predicting the reference value at all, and the operator could equally well report the mean of the original data. The SEP should lower than SD and ideally the ratio of RPD should be 5 or higher, but at least 3 that indicate efficient NIR reflectance prediction. (Williams, 2001)

$$\text{RPD} = \frac{\text{SD of val set}}{\text{SEP}} \quad (20)$$

8. Application of Near Infrared Spectroscopy for Wood Industry

Easty *et al.*(1990) used reflectance NIR to quantify hardwood-softwood ratios in paperboard, and to estimate lignin content in hardwood and softwoods. They found that separate calibrations were required for two types of woods. Using second derivative spectra, they identified wavelengths which correlated with lignin content at 1680 nm and in the 2100 to 2200 nm region.

Schultz and Burns (1990) compared NIR and FTIR techniques for analysis of hardwood and softwood for lignin, hemicellulose, and cellulose content. The NIR technique was judged to be superior for this analysis on the basis of simplicity of sample preparation, shorter scanning time, and smaller standard errors. The best results were for NIR analysis of lignin, with standard errors of calibration less than 1% covers the range of 10 to 30% in pine and sweetgum.

Wright (1990) reported that NIRs were able to predict pulp yield and cellulose content in wood powder samples. Using a filter instrument employing five wavelengths, they indicated a standard error of about 2%.

Wallback *et. al.* (1991) compared results for the analysis of birch pulp using NMR FTIR, and NIR. The pulping process is essentially a delignification reaction in which lignin's phenylpropane units are cleaved at ether linkage with an alkaline reagent. NIR bands from cellulose, hemicellulose, and lignin were overlapped, but it could be resolved by using partial least square (PLS) analysis. They developed PLS models for Klason lignin, glucose, and xylose. NIR showed the best prediction results with a standard error of 1% or less for all three components. The range of measurements was xylose: 21-26%; glucose: 55-57%; and Klason lignin: 0-25%.

Michell (1995) had investigated NIR spectroscopic technique to soda charge at Kappa number 15, total lignin, hot water, and alkali soluble of pulp yield. The woods were drawn from three trees in 10 provenances of *E.globulus*. The calibrations were

compared between standard regression, Norris regression, PLS from NSAS program and PLS from Unscramble program. The conclusion reported for the regression models built by original spectra, the results obtained by standard, Norris and two form of PLS regressions had similar precision. When the second derivative form of the spectra were used to develop the model, the PLS regression showed more precise results. Results obtained by using the second derivative forms were not always more precise than those obtained by using the normal form, but would be expected to be so in cases where particle size varied substantially.

Marklund *et al.* (1999) studied the relationship between the softwood raw material spectra and the properties of the kraft pulps, expressed in terms of physical parameters for pulps and strength properties for the corresponding hand sheets. The kraft pulps were made from 20 different types of wood samples. The relationship between the spectra of wood chips and fully bleached pulp, and properties were used to generate predictive models for fibre properties and strength parameters. Orthogonal signal correction (OSC) was treated the spectra, it was shown that the NIR spectra of milled wood chips have nearly the same predictive abilities as those of the bleached pulps.

Reiter *et al.* (1999) used the FT-NIR spectroscopy for characterization of pulp and chemical additives. An identity check of alternative products of the same raw material group can be quickly determined. For the control of product quality a reference library is needed. Quantitative analyses require and increased calibration effort.

Meghanathan *et al.* (2000) investigated the modeling of Kraft pulping process base on on-line data from NIR analyzer. It showed the application of an in-situ NIR analyzer to develop a kinetic model for the batch kraft pulping process. The NIR analyzer provides measurement of circulating liquor stream at one-minute intervals. Base on these online data, the instrument is calibrated using the PLS model to predict the effective alkali, sulfide, lignin, and carbohydrate contents. The major thrust of this paper estimated the pulp yield and kappa no. from the instrument predictions using the batch data. A preliminary kinetic model of the batch kraft process is then developed using the liquor properties predicted by the PLS models. Model predictions used in the lab.

Several researches were carried out in wood by using NIRs for classification (Schimleck *et al.*, 1996) they succeed on the classification of *E.globules*, *E.nitens*, *E.grandis* and *radiate Pine* on Principle Component Analysis(PCA). In 1997 they estimated the chemical compositions of *E.globules* and *E.nitens*. PLS models were developed using the second derivative spectra, and these models that could be used to estimate carbohydrate component. The models predicted the *E.globules* yielded more precise result than the one for the *E.nitens* wood. In addition, PLS models based on glucan and xylan contents showed an improved fit relate to simple regression on glucan content, for the pulp yield data of *E.nitens* wood, but not for *E.globules* woods.

In addition, Schimleck and Yazaki (2003) studied *Pinus radiate D. Don* bark properties in term of water extractive, NaOH extractive and stiasny value. The

calibrations developed for Stiasny value and hot water extractive were successful in predicting these properties in separate test set. The calibration developed for the Stiasny value could have practical importance as the chemical method used to determine the Stiasny value particularly is time consuming. However, the NaOH extractive calibration did not indicated good result. In 2004, Schimleck and Evans studied the possibility of developing calibrations for several tracheid morphological characteristics using NIR spectra obtained section from the radial-longitudinal face of wooden strips cut from *P. radiate* increment cores. That were coarseness, perimeter, radial, tangential diameter and wall thickness. The results showed that only the calibrations developed for trachied coarseness, wall thickness and tangential diameter performed well when tested.

Ali *et al.* (2001) studied in the topic of spectroscopic studies of the aging of cellulose paper. FTIR and NIR spectroscopy had been used to characterize the aging cellulose paper. The objective was a technique development to evaluate the condition of paper insulation in electrical transformers. There were different in the spectra of new and aged materials from the different sources, which provided the basis of a “fingerprint” method to categories papers into different “families”. Chemometrics methods was applied to NIR spectra of age cellulose that performed a correlation of 0.994 between the spectra and time of ageing, with an error of prediction of 95 h for samples up to 3000 h of ageing.

Rodrigues *et al.*(2001) used Furiour transform infrared spectroscopy (FT-IR) to estimate monosaccharide content in *E.globulus*; glucose, xylose, galactose, mannose and rhamnose. All 38 samples with 9 years old tree from provenance trials were used in this study. The multivariate analysis found better results than the univariate method. That indicated with increase R^2 , lower SEC and SEP error, although the univariate method also provided good fits. From the result, the FTIR techniques could be used in large scale breeding programmes to measure wood monosaccharide composition with inexpensive effort and in shorter time than wet-lab methods.

Tsuchikawa *et al.* (2003) studied a nondestructive discriminant analysis on the basis of combination of NIR spectra wood and Mahalanobis’ generalized distance. The distances between softwoods were relatively close independent of analytical pattern, so it was difficult to explain the difference in their spectroscopic characteristics. On the other hand, the distances between hardwoods were large, that indicated to easy classified.

Xiaomei *et al.* (2003) researched in the topic of application of near-infrared spectroscopy in pulp and paper industry. The applied topics were the evaluation of waste paper deinking efficiency, the online measurement of kappa no. during pulping, the kappa no. of the pulp, the lignin content of the pulp, and the moisture content of the coating are introduced as examples.

Kelly *et al.* (2004) used near infrared spectroscopy to predict the chemical and mechanical properties of solid loblolly pine wood by PLS model. The chemical properties were extractive, lignin, and carbohydrate. The mechanical properties were

stiffness, and strength. The correlations were very strong, with the correlation coefficients generally above 0.80. The mechanical properties could also be predicted using the spectra range 650-1150 nm.

Wenhao *et al.* (2004) studies the on-line pulp Kappa no. determination with spectroscopy during batch pulping. They predicted the lignin content in pulp through the measurement of dissolved lignin in cooking liquor. It was discovered that the absorption of red liquor from masson's pine in 460-580nm wave range resulted mainly from lignin-sulfonate with high molecular weight, where the absorption of black liquor from mass's pine in 800-900nm wave range resulted mainly from Kraft lignin, and Kappa no. maintained a linear relationship. Mathematic model of online Kappa no. determination with visible spectroscopy and NIR were developed by chemometrics for bath sulfite and Kraft cooking. It was obviously that with this method to control sulfite pulp Kappa no. and predicted the end point of pulping during cooking process, the first grade ratio of pulp had been improved from 55% to 70%, the pulp Kappa no. could be controlled within ± 2 units.

Figure 25 illustrated absorbance of several paper specimens in the near-infrared region.

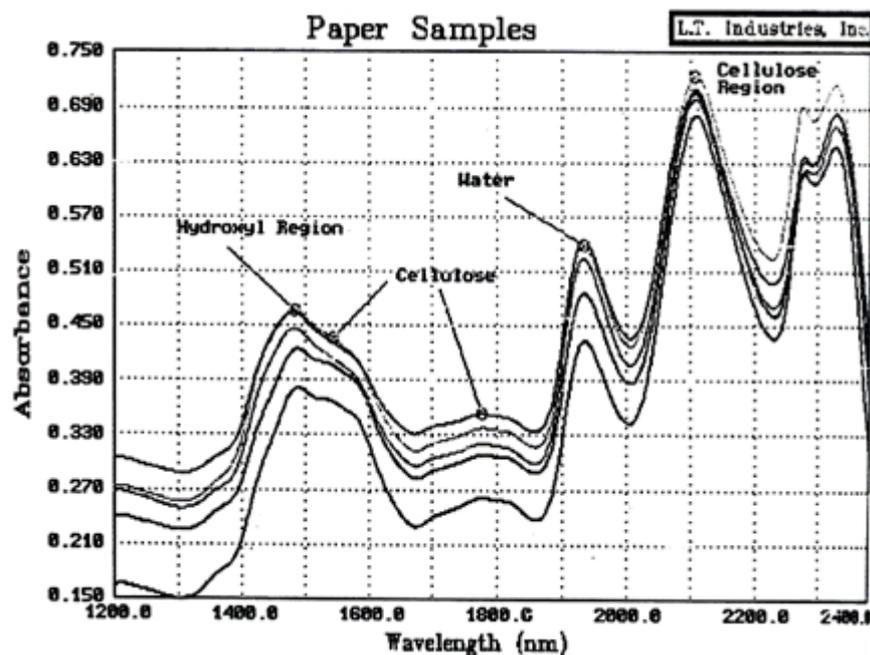


Figure 25 Typical absorbance spectra of paper, showing absorbance regions attributed to cellulose and water.

Source : Connors(1995)

9. Advancetage of NIR

NIR has many advantages compared to chemical and other instrumental methods of wood and pulps analysis, although the results from other technique are more accurate than those from the NIR methods.

1. The NIR method is nondestructive.
2. This method is more rapid than other analytical methods.
3. Very little sample preparation is required for NIR measurements.
4. There is no need of using a large amount of chemicals.
5. This is the low cost analysis in long term consideration.

From the literature review the NIR method could be replaced the conventional method in the pulp and paper industry. So, the pulp and paper industry in Thailand should be rapidly developed for the advancement, safe for environment, and reduce cost for competitive in the future.

MATERIAL AND METHODS

1. Materials

1. 1-Methylimidazole (Merck, Germany)
2. Acetic acid (J.T. Baker, USA.)
3. Acetic anhydride (Merck, Germany)
4. Ammonium hydroxide (28-30%, J.T.Baker, USA.)
5. Ethanol (99%,Merck, Germany)
6. D-Galactose (Fuka, Switzerland)
7. D-Glucose-monohydrate(Merck, Germany)
8. D-Mannose (Fuka, Switzerland)
9. D-Xylose (Fuka, Switzerland)
10. Dichromethane (Carlo Erba, Germany)
11. *Eucalyptus camaldulensis* wood powder, from Siam Pulp and Paper Public Co., Ltd. 70 samples
12. Ferric chloride(Fuka, Switzerland)
13. Sulfuric acid (72%, J.T. Baker, USA.)
14. Inositol (Fuka, Switzerland)
15. L-Arabinose (Koch-light, Ltd., UK.)
16. Oricine monohydrate (Fuka, Switzerland)
17. Sodium chlorite (APS, Australia)
18. Sodium borohydride (Merck, Germany)
19. Toluene (Fisher Scientific, UK.)

2. Equipments

1. Balance 4 digit (Precisa, 120A, USA.)
2. Pulping unit (6-Batch Digester, L&W, FI-229, Finland)
3. Screening (Somerville screen /size 0.15 mm, BUCHEL, BK-34, Natherland)
4. Sieve 425 μm (D-42759, 40 mesh, Retsch, Germany)
5. Sieve 250 μm (60 mesh, Endocoris, England)
6. Carbolite Furnace (CWF 11/23/201, England)
7. Gas Chromatography Instrument (Agilent Technique, 6890N, USA.)
8. Hot air oven (Binder, Germany)
9. Hot plate (Barndstead Electromal, EME6 0250/CEB, UK.)
10. Near Infrared Reflectance Spectrophotometer (BRAN+LUEBBE, InfraAlyzer 500, Norderstedt, Germany)
11. Soxhlet Extraction apparatus (Buchi, B-811, Switzerland)
12. UV-spectrophotometer (Japan spectroscopic, 7800, Japan)
13. Water bath (Memmert, WB14, Germany)

3. Methods

3.1 *Eucalyptus* Sample Preparation

70 samples of *Eucalyptus camaldulensis* wood chip were obtained from The Siam Pulp and Paper Public Company Limited, Thailand. They were inspected for defects of different particle size. Therefore, the wood chips were ground with Wiley mill. After that, it is screened through the size 40-60 mesh. All wood powder that remains on 60 mesh of screen was controlled moisture by drying in oven at 50°C for two days, and it is cooled in desiccators. After that it was kept in a plastic bag to store the samples.



Figure 26 Presented the method of control moisture content in wood powder by drying in oven at 50°C.

3.2 Spectral Reflectance Measurements

Wood powder was placed centrally upon the cup holder, and it was irradiated from the top by the light source. Wood powder were scanned from 1100-2500 nm with Infra-Alyzer 500 spectrometer (Bran+Luebbe, Germany), and the data were collected every 2 nm (about 701 points for each spectrum). The wood powder was scanned at the temperature of 20-25°C. Thirty repetitive scans were accumulated in the computer memory. The scans were averaged and transformed to $\log(1/R)$, where R is reflectance, and then stored in computer files. The scanning of wood sample was performed at ambient temperature, and the spectral data for 3 packs was recorded by microcomputer with the SESAME software program.



Figure 27 Closed cup packing (a) and sample holder using in the near infrared spectrometer Infra-Alyzer 500 (Bran+Luebbe) (b).

3.3 Chemical Measurement

3.3.1 Analysis of Total Pulp Yield

100g of dried wood chip were pulping with the solution of NaOH for 2 hours at 170°C to receive pulp yield. The total percent of pulp yield was calculated by dividing the dried pulp weight with dried wood chip weight, and multiplied by 100.

3.3.2 Analysis of Screen Pulp Yield

The pulp obtained after pulping was filtered to the screening process. The residue pulp on the screen is the reject. The pulp that passed through the screen was washed with water, then dried and weighted. Percent screen pulp yield was calculated by dividing the dried residue pulp weight with the dried wood chip weight, and multiplied by 100. This analysis followed the method from the handbook of pulping and papermaking. (Biermann 1996)

3.3.3 Analysis of Wood Extractive

Of the wood powder, at least 4 g were required for determination in duplicate wood powder was weighted into the extraction thimble in the Soxhlet apparatus. It was extracted with 150 ml of solvent that contain the mixture of ethanol (1 volume) and toluene (2 volumes). The extraction apparatus was heat to boiling temperature of solvent. The cycle rate of the solvents was used at least 6 times per hour over 4-5 hour period. The flask was removed from the apparatus, and it was evaporated the solvent until dryness. The extractive was dried in oven for one hour at 100±5°C. The residual was weight and calculated by divided with the weigh of wood powder, and multiplied by 100.

3.3.4 Sample Preparation for Chemical Analysis

About 10 g of wood powder was extracted with Ethanol-Toluene in Soxhlet extraction apparatus after that, the wood powder was transferred to beaker. The excess solvent was removed by washing with 95% ethanol for 4 hours. Then, the excess solvent was removed with suction and the sample was further washed with hot water for 1 hour. The wood powder was filtered and it dried at room temperature after that it was stored in an air-tight container for further chemical analysis. The moisture content in wood was calculated.

3.3.5 Analysis of Holocellulose

The extractive-free wood powder was weighed about 5 g into an Erlenmeyer flask. The 160 ml of distilled water, 0.5 ml of glacial acetic acid and 1.5 ± 0.1 g of sodium chlorite were added, respectively. This flask was placed on a steam bath, of which the temperature is adjusted about 70-80°C. The flask was heated for 1 hour at the reaction temperature, and the contents were mixed by occasional swirling. Then, without cooling, adding 0.5 ml glacial acetic acid followed by 1.5 g sodium chlorite, respectively. The heating was continuous for an additional hour. At the end of second and third hours, the additions of acetic acid and sodium chlorite were repeated. At the end of the fourth hour, the flask was placed in an ice bath until the contents have cooled below 10°C. The holocellulose was filtered on glass crucible pore No.1 and washed with cool water to remove the color and odor of chlorine dioxide, after that it was dried in the oven at $100 \pm 5^\circ\text{C}$ for 24 hours. Move to a desiccator, and let it sit 1 hour and weigh. Calculate the holocellulose content by dividing with weight of wood powder and multiplied with 100.

3.3.6 Analysis of α -Cellulose

Weigh out about 1.5 ± 0.1 g of holocellulose into beaker, and added 75.0 ml of 17.5% NaOH. The temperature was controlled in 25°C. The holocellulose was stirred with the stirrer until it was completely dispersed. When the holocellulose was dispersed, the stirrer was rinsed for removing the adhered holocellulose with 25 ml of 17.5% NaOH reagents again. After that, the suspension was stirred with a glass rod and maintained temperature in a water bath at 25°C. After 30 min from the first addition of the NaOH reagent, 100 ml of distilled water at 25°C was added to the suspension and stirred thoroughly with a glass rod. The beaker was stood in a water bath for 30 min so that the total extraction time was 60 ± 5 min. The residual was filtered with a glass filtering crucible pore no.3, after that washing with distilled water until the washing solution was neutral. Pour 40 ml of 10% acetic acid into the crucible was added and held for 5 min, from the time the suction was released, place the cellulose overnight in the oven $100 \pm 5^\circ\text{C}$. Move to a desiccator and let it sit for 1 hour and weighed. Calculate the α -cellulose content by dividing with the weight of holocellulose and multiplied with 100.

3.3.7 Analysis of Lignin Content

Weighed 1 ± 0.1 g of extractive free wood powder into beaker and hydrolyzed with 15 ml of 72% H_2SO_4 at 2°C in cooling bath for 1 hour. The solution was kept continuously hydrolyzed at 20°C for 2 hours to ensure completely reaction. The solution was diluted to 3% H_2SO_4 with deionised water that make up the total volume of solution to 575 ml. The solution was refluxed for 4 hours. The refluxed solution was stood over night and filtered through a-glass filter crucible no.4. The residue was washed with hot water and dried overnight at $100\pm 5^\circ\text{C}$ in oven. Move to a desiccator and let it sit 1 hour and weighed by divided with the weight of wood powder and multiplied with 100.

3.3.8 Analysis of Pentosan Content

Weighed 0.2-0.3 g of extractive free wood powder into a boiling flask and added 20 g of NaCl and 100 ml of 3.85 N HCl into the flask. The flask was connected to distillation apparatus, followed by heating the solution. While maintaining the solution to a constant volume of 100 ml in the boiling flask by adding HCL drop wise from a separation funnel which was connected to the boiling flask. The solution was heated until the distillate volume reached 225 ± 10 ml, then the distillation was stopped. The distillate was adjusted to 250 ml with deionized water. Pipeted 10 ml of the distillate to formed complex solution with 25 ml of orcinol reagent. The preparation of orcinol reagent, 0.4 g of orcinol and 0.5 g of ferric chloride were dissolved in 1000 ml of 11 N hydrochloric acid. The complex solution was stood at 25°C for 60 min and measured the absorbance by UV-VIS spectrophotometer. The absorbance was converted to concentration of pentosan by comparing with the standard graph of xylose. Pentosan content was calculated by divided with the weight of wood powder and multiplied by 100.

3.3.9 Analysis of Ash Content

The empty crucible was carefully cleaned and ignited in a muffle furnace at $525\pm 25^\circ\text{C}$ for 30-60 min. After ignition, the crucible was cooled slightly by placed in desiccators. This crucible was cooled to room temperature, and weighed. The wood powder was weighed at least 1 g and put into the crucible. The crucible was placed in a muffle furnace at temperature about 100°C . Slowly, raised the temperature to 525°C , the ignition time about 6 hour, so that the sample becomes carbonize without flaming. When the specimen was completely combusted, the crucible was removed from the muffle furnace. Then it was placed in the desiccators and cooled to room temperature. Reweigh the crucible with ash and calculate the ash content by divided with the weight of wood powder and multiple with 100.

3.3.10 Analysis of 1%NaOH solubility

Weigh 2.0 ± 0.1 of wood powder into beaker. Add 100 ml of 1%NaOH solution and stirred with glass rod. The beaker was cover with watch glass and placed in a water bath to maintain the temperature at 97 to 100°C for 60 min. The

wood powder was stirred with the glass rod for 5, 10, 15, and 25 min after placing in the water bath. At the end of 60 min, the wood powder was filter to a filter glass crucible and washed with 100 ml of hot water. The 10% acetic acid (25 ml) was added and allowed to soak for 1 min before suction. This step was repeated two times. Finally, the wood powder was washed with hot water until acid free. The crucible and wood powder was dried overnight in the oven at $105\pm 3^{\circ}\text{C}$ to constant weight, cooled in a desiccators and weighed.

3.3.11 Analysis of Monosaccharide

Procedure

Acid Hydrolysis : Weighed 0.35 g of extractive free wood into 150-ml beakers. Add 3ml of 72% sulfuric acid with pipette into a beaker. The contents in the beaker was stirred with glass rod until the specimen starts to dissolve. The beaker was placed in a $30\pm 0.5^{\circ}\text{C}$ water bath for 11 hour. The specimens was stirred every 5-10 min. At the end of the hour, 84 ml of distilled water was added, and refluxed for 4 hours. 10 ml of myo-inositol internal standard solution was added in to flask. After mixing, the solution was cooled. The lignin residue was filtered through a glass crucible pore No.4. The filtrated was adjusted volume to 140 ml. 11 ml of 28-30% ammonium hydroxide solutions was added to neutralize solution.

Reduction: A 2-ml aliquot was pipeted into a 125-ml Erlenmeyer flask. 35 ± 1 mg of sodium borohydride was added, the flask was placed in a 40°C water bath for 90 min. The content was stirred occasionally. After reduction, the excess of sodium borohydride was decomposed by adding glacial acetic acid drop by drop at a time.

Acetylation and Extraction: After cooling the solution to ambient temperature, 2 ml of 1-methylimidazole and a stirring bar were added, immediately, 20 ml of acetic anhydride was followed for monosaccharide reduction. The mixture was continuously agitated with a magnetic stirrer at room temperature for 20 min. The heat generated has been promoted the acetylating to complete reduction. After 20 min, 30 g of crushed ice and 70 g distilled water were added into the flask, continue to stir for at least 20 min. The mixture to 250-ml was transferred to separation funnel and extracted successively with 10-,5-, and 5-ml portions of dichloromethane. The over all of extracted dichloromethane was combined in beaker. the beaker was putted in a well-ventilated fume hood for at least 1 hour. The 2 ml of dichloromethane was added to the dried alditol acetate mixture. 2 μl of the mixture was injected into GC.

GC Condition

Equipment : Agilent Technologies
 Column : BPX-70(70% Cyanipropyl polysilphenylene-siloxane) capillary column (SEG compny) , 25m x 0.32mm ID, 0.25 μm film thickness
 Carrier gas : He
 Flow rate : 2 ml/min
 Pressure : 13.59 psi
 Detector : FID (Flame ionization detector) at 300°C

- Gradient : Oven temperature was set at 150°C, and then raised at 10°C/min to 190°C where it was held for 2 min. After that the temperature was raised at 5°C/min to 250°C, where it was held for 10 min.
- Injection : Standard and samples (2 µl) were injected using the split mode ratio 50:1, injection port held at 260°C

Determination of Accuracy

The pure cellulose was hydrolyzed with 72% H_2SO_4 . Then the followed step was reduced, acetylated, and extracted along with 3.3.11.1. Finally, the mixture was injected into GC.

3.4 Data Analysis

The SESAME software (ver. 3.0, Bran+Luebbe, Germany) was employed for the spectral data collection. The spectral data obtained were converted into files for the NSAS (Fross NIRSystems, Silver Spring, U.S.A.) and Unscrambler (ver.7.08: CAMO AS, Trondheim, Norway) software program. The purpose of calibration model was to relate the concentration of some chemical component analyze found in sample to the spectral data collected from that sample. There were several programs and statistics for computation.

3.4.1 Multiple Linear Regressions (MLR) by NSAS Software Package

The basis for an MLR calibration consist of two sets of data: the first is the set of reflectance, or transmittance values, measured at the different wavelengths for the calibration set samples, and the second is the set of the reference property values for the corresponding samples. If a transformation is selected for the calibration, for exsample Absorbance or Second Derivative, the transformed values will be used instead of the measured ones.

In MLR calibration, the Step-up search is a faster way to determine a combination of wavelengths than the Combination search. It is especially suitable for scanning instruments, which provide a large number of wavelengths, but it would be very time-consuming to determine all wavelengths by testing all possible combinations.

3.4.2 Partial Least Square Regression (PLSR) by Unscramble Software Package

All the spectra were pretreated by the different transformation as Multiplicative Scattering Correction (MSC), Savitzky-Golay second derivative, and both of these transformations. In the Moving Window Partial Least Square Regression (MWPLSR) calculations, the window size was 20 spectral points and the mean centered spectra in the region of 1100-2500nm were applied. Ten residue lines corresponding to one to ten PLS component numbers were plotted as a function of the position of the window. All these calculations were carried out by in-house-written programs in MATLAB (ver.5.3: The MathWorks, U.S.A). PLS1 (Unscrambler software) was applied to the spectral region that was selected by MWPLSR to

develop the multivariate calibration and prediction models that predicted the concentrations simultaneously. The optimum number of PLS factors was selected by considering the factor number at which the lowest standard error of prediction (SEP) was obtained and increased to the next number. The comparison of the PLS results were obtained using the regions selected by MWPLSR and the whole region.

RESULTS AND DISCUSSION

1. Raw Material

A total of 70 samples of *Eucalyptus camaldulensis* wood powder were obtained from The Siam Pulp and Paper Public Company Limited, Thailand. The general property of milled wood was presented in Figure 28. There were inspected for the different particle sizes. By separating the sizes with sieve on the size 40-60 mesh at the factory. This size of wood powder was used for NIR and chemical analysis.



Figure 28 *Eucalyptus camaldulensis* wood powder sample.

2. Chemical Analysis

2.1 The Pulp Yield Property Value

The total pulp yield and the screen pulp yield value of soda process obtained from Siam Pulp and Paper Public Company Limited, Thailand. The results of pulp yield value of 70 samples were shown the range of minimum and maximum value in Table 2.

Table 2 Summaries of statistic for the measurements of pulp yield property

| Parameter | %Total pulp yield* | %Screen pulp yield* |
|-----------|--------------------|---------------------|
| Min | 38.02 | 37.94 |
| Max | 50.63 | 50.43 |
| Mean | 45.97 | 45.88 |
| SD | 2.89 | 2.91 |

* Unit (%W/W)

The total pulp yield values measured from pulp yield that obtained after Soda pulping. These were in the range of 38.02% to 50.63% (w/w). The screen pulp yields were in the range of 37.94 to 50.43%. The correlation between two values was 0.999 (Figure 29). This means, both values were related.

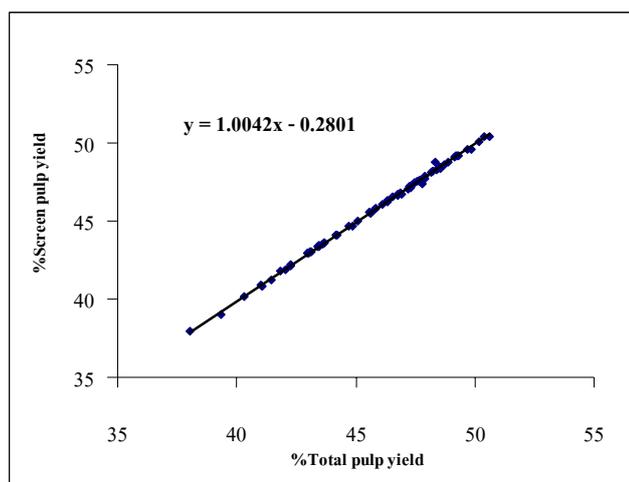


Figure 29 Relationship between total pulp yield and screen pulp yield

2.2 Chemical Composition Values

The main compositions of wood were holocellulose, α -cellulose, and Lignin, on the other hand, the minor components were ash and wood extractive. In this experiment the value of chemical component (lignin, pentosan, extractive, ash, α -cellulose, holocellulose, and 1% NaOH solubility) were supplied from the Siam Pulp and Paper Public Company Limited, Thailand. The summaries of chemical compositions in the range of minimum and maximum value of 70 samples were presented in Table 3.

Table 3 Summaries of the chemical composition of wood powder.

| Parameter | % Holo-cellulose* | % α -cellulose | % Lignin* | % Pentosan | 1%NaOH solubility | % Extractive* | %Ash* |
|-----------|-------------------|-----------------------|-----------|------------|-------------------|---------------|-------|
| Min | 73.82 | 39.14 | 25.42 | 12.64 | 13.71 | 0.80 | 0.61 |
| Max | 83.06 | 48.93 | 32.12 | 17.86 | 26.80 | 4.31 | 1.63 |
| Mean | 79.28 | 43.76 | 28.39 | 15.38 | 18.36 | 2.44 | 0.84 |
| SD | 2.17 | 2.22 | 1.14 | 1.21 | 2.94 | 0.94 | 0.14 |

*The percentage combination of holocellulose, lignin, pentosan, extractive, and ash about 100%

2.3 Monosaccharide Value

In this experiment, the monosaccharide determination was determined by GC method. The wood powder was hydrolyzed and acetylated of individual monosaccharide before the injection.

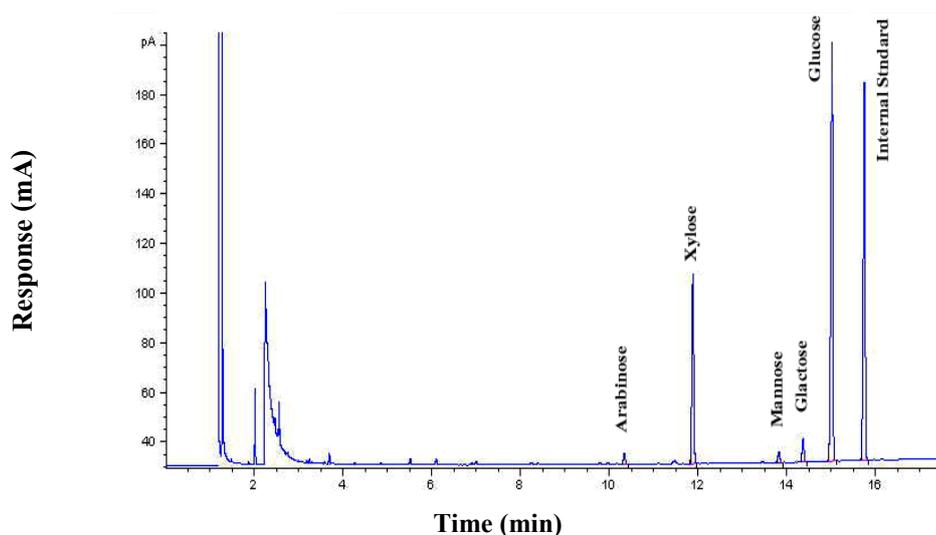


Figure 30 GC chromatogram of alditol acetate derivative of monosaccharide in *Eucalyptus* wood powder.

Gas chromatography (GC) was powerful technique for analysis of monosaccharide. However this technique was only suitable for volatile samples, free monosaccharide could not be analyzed by gas chromatographic directly because of their low volatility. The alditol acetate derivatisation was widely used for determining the composition of monosaccharide mixture. The inositol was used to calculate as internal standard method. The chromatogram was shown in Figure 30. The monosaccharides found in wood powder were glucose and xylose. In addition, there were few of arabinose, mannose, galactose content. The summaries of the range of minimum and maximum content of monosaccharide in wood powder were presented in Table 4.

Table 4 Summary of the carbohydrate component value of wood powder.

| Parameter | %Glucose | %Xylose | %Arabinose | %Galactose | %Mannose |
|-----------|----------|---------|------------|------------|----------|
| Min | 43.74 | 8.81 | 0.38 | 1.16 | 0.28 |
| Max | 52.69 | 16.57 | 1.01 | 3.67 | 1.56 |
| Mean | 48.30 | 12.94 | 0.61 | 2.19 | 0.67 |
| SD | 1.80 | 2.29 | 0.16 | 0.51 | 0.21 |

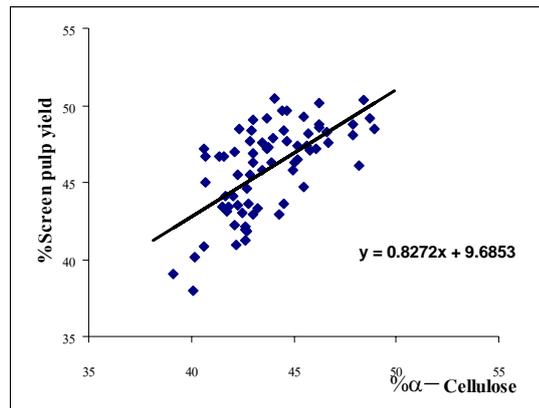
The accuracy test of gas chromatography method is presented in Table 5. The pure cellulose was used instead of the wood powdered and the test procedures were the same.

Table 5 Present recovery of glucose that received in pure cellulose.

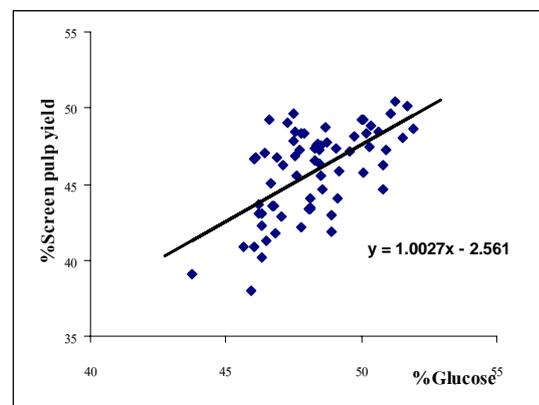
| No. | Recovery (%) |
|-----|--------------|
| 1 | 100.40 |
| 2 | 98.67 |

2.4 Comparison between The Chemical Value

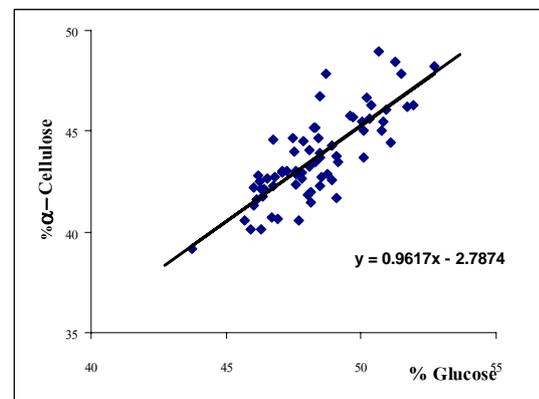
Theoretically the main composition paper was cellulose, so the pulp yield depended on the percentage of cellulose content. And the cellulose was contained the glucose component. From the conventional method, the correlation between the percentage of pulp yield value and α -cellulose was 0.633. The correlation between the screen pulp yield value and the glucose content was 0.603, and between the α -cellulose content and the glucose content was 0.801, all data illustrated in Figure 31.



(a)



(b)



(c)

Figure 31 The correlations between the screen pulp yield value and α -cellulose content (a), the screen pulp yield and glucose content (b), α -cellulose content and glucose content (c).

3. NIR Analysis

The NIR spectral data were obtained in the range of (1100-2500 nm). The resolution of the adjacent wavelengths was taken as 2 nm. The total number of absorbance points was 701 points. In NIR reflectance spectroscopy the baseline slopes upwards with increasing wavelength due to the sample varies with particle size (Michell, 1995). The derivative method is powerful in separating superimposed bands and correcting for baseline slope (Ozaki, 2000). The peak signal of an averaged original NIR and its second derivatives spectrum of *Eucalyptus camaldulensis* wood powder were presented in Figure 32.

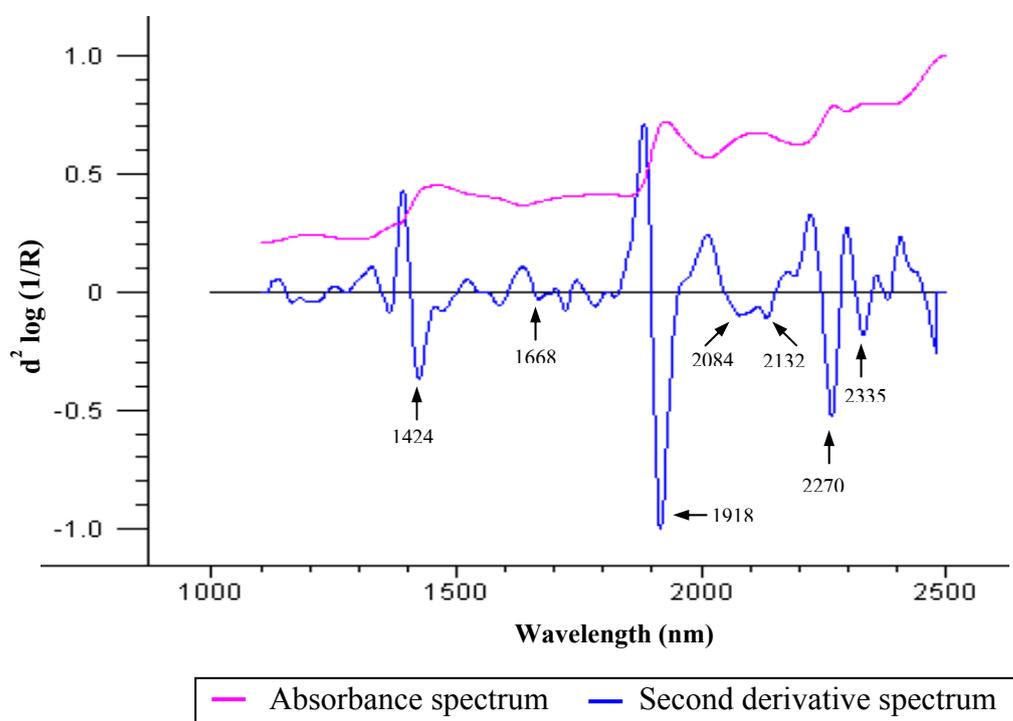


Figure 32 An absorbance NIR and its second derivative spectrum of *Eucalyptus camaldulensis* wood powder sample.

From Figure 32, both regions in 1600-1800 nm and 2000-2500 nm presented large variation. The bands in the region of 1600-1800 nm rose from the first overtone, and the bands in the region of 2000-2500 nm mainly rose from C-H, N-H, O-H combination band. (Shenk *et al.* 1992; Osborne *et al.* 1993)

The other peaks have been reported from the previous paper. Two peaks in range of 1350-1450 nm specifically at 1424 nm and 1848-1968 nm specifically at 1918 nm, observed due to high absorbance of water molecule. Bands near 2084 nm rose from OH stretching combined with OH, and CH deformation band of xylan. Bands near 2132 nm rose from C-H stretching of lignin and wood extractives (Schimleck, 1997). The region of 1668 assigned as aromatic structure of lignin (Siesler, 2002). Band at 2270 nm assigned as O-H stretching combined the region of

C-O stretching of cellulose. Moreover, the band in 2335 nm assigned as C-H stretching combined CH₂ deformation of glucose based polymers such as cellulose (Burns, 2001).

3.1 Spectral Measurements

Figure 33 showed the original spectra of *Eucalyptus camaldulensis* wood powder in the region of 1100-2500 nm. Those spectra present the scattering effects of particle size. The 70 average spectra for those repeating measure the same sample illustrated in Figure 33.

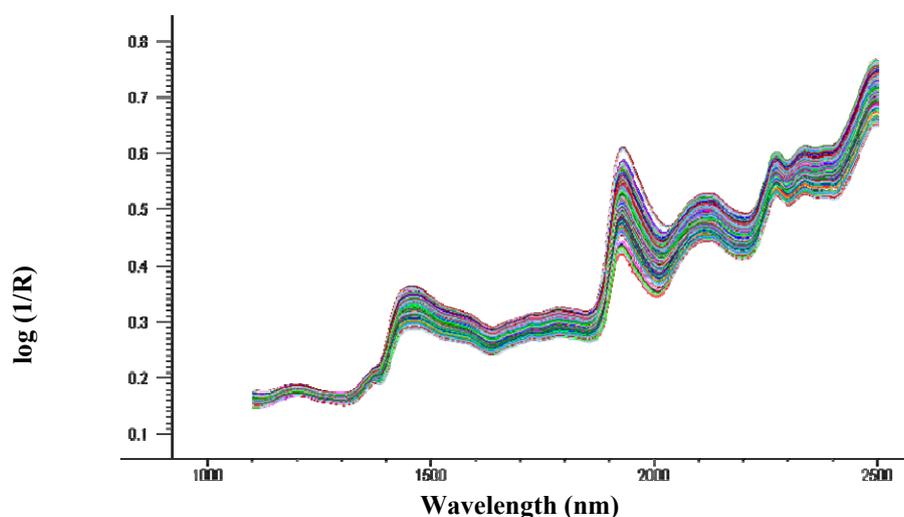


Figure 33 Seventy original near infrared spectra of *Eucalyptus camaldulensis* wood powder.

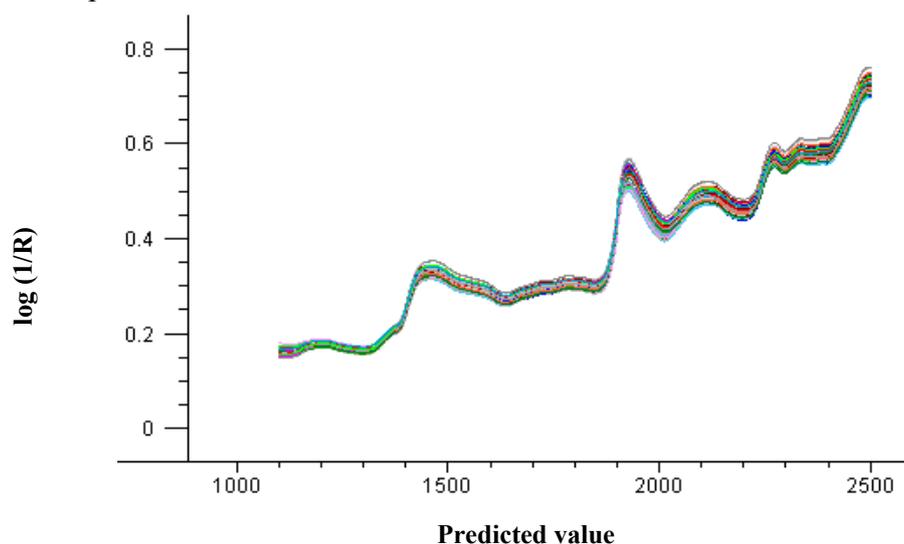


Figure 34 Averaged raw near infrared spectra of 70 *Eucalyptus camaldulensis* wood powder.

Calibration of spectral volumes was not as straightforward as standard calibration procedures where the chemical component of interest was measured on a homogeneous sample so that a single spectrum related directly to a corresponding measurement.

3.2 Pretreatments Data

In this experiment, the use of Multiplicative Scatter Correction (MSC), second derivative and the combination of these pretreatment was performed before developing calibration.

MSC was specific transformation for spectra. It consisted in fitting a separate spectra regression line to each sample spectrum that expressed as a function of the average value for each wavelength. MSC correct in scattering with particle size. The MSC pretreated spectra were shown in Figure 35.

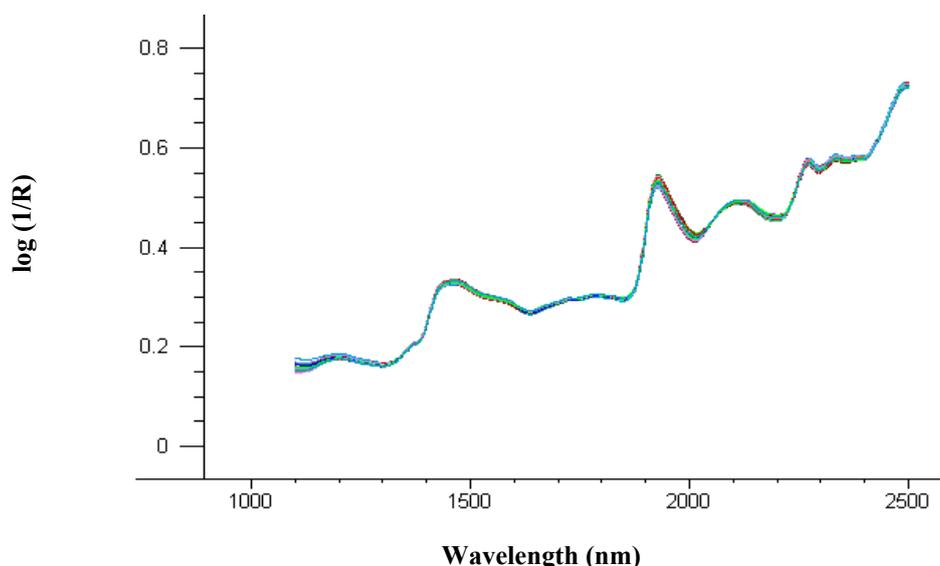


Figure 35 MSC pretreated spectra of 70 *Eucalyptus camaldulensis* wood powder.

The Savitzky-Golay derivative was a practical alternative to simple difference and gap derivatives. The Savitzky-Golay algorithm was based on fitting a polynomial to a range of points around each data point in the spectrum. The Savitzky-Golay second derivative spectra of 70 *Eucalyptus camaldulensis* wood powder are shown in Figure 36.

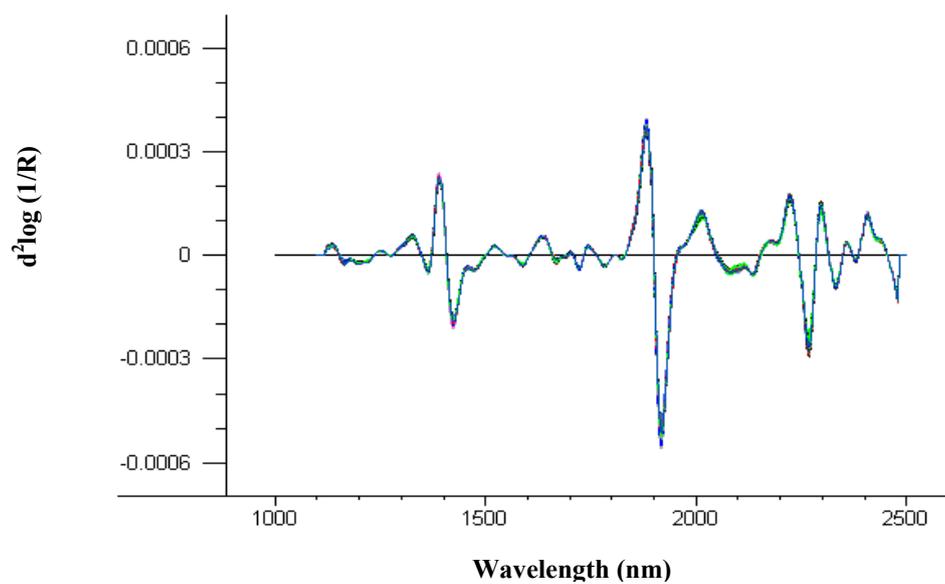


Figure 36 Savitzky-Golay 2nd derivative spectra of 70 *Eucalyptus camaldulensis* wood powder.

The results of Savitzky-Golay second derivative of chemical component were showed in Figures 37, 38, 39, and 40. The tentative assignments of NIR bands were shown in Table 6 and Table 7.

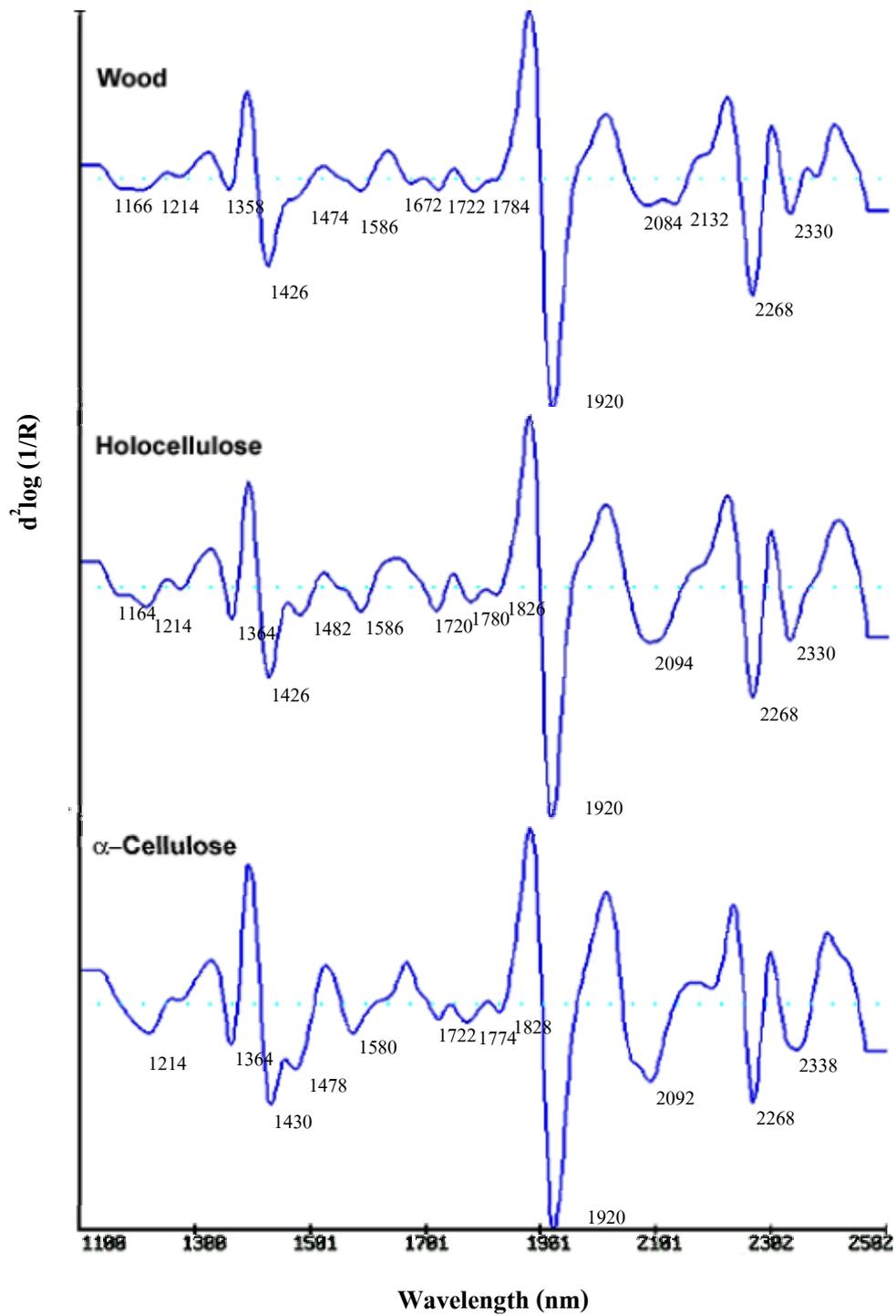


Figure 37 Second derivative spectra of *Eucalyptus camaldulensis* wood powder, holocellulose, and α -cellulose content.

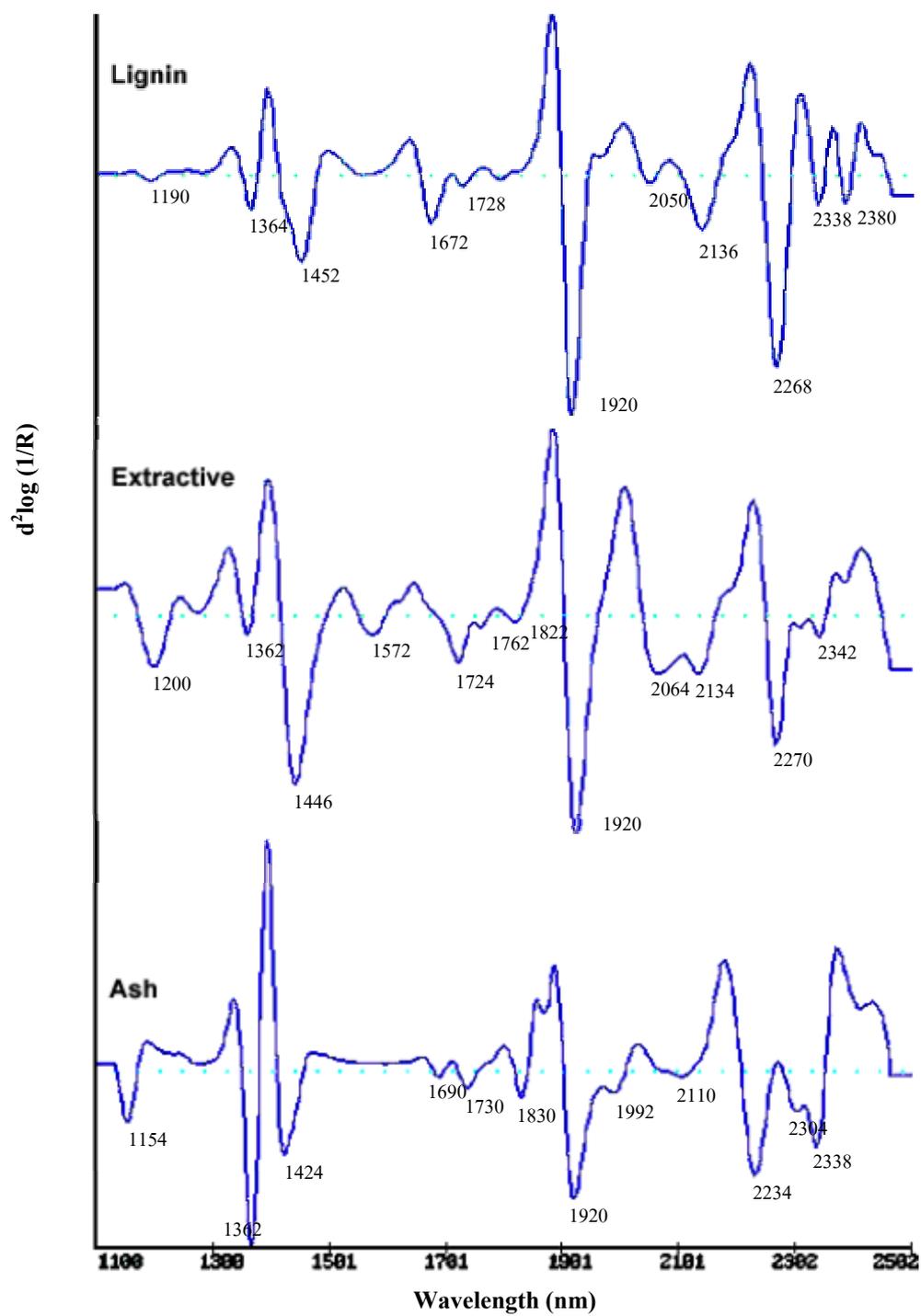


Figure 38 Second derivative spectra of lignin, wood extractive, ash from wood powder.

Table 6 Tentative assignments of NIR bands for *E. californica* woods mill and its components (Burn,2001, Ozaki, 2001, Osborn,1997, Shimleck,1997)

| NIR band (nm) | | | | | | Assignment | Structure |
|---------------|---------------|---------------------|--------|------------|------|---|--------------------------|
| Wood | Holocellulose | α -Cellulose | Lignin | Extractive | Ash | | |
| 1166 | 1164 | | 1190 | | 1154 | C-H stretch 2 nd overtone | CH ₃ |
| 1204 | 1214 | 1214 | | 1200 | | C-H stretch 2 nd overtone | CH |
| 1358 | 1364 | 1364 | 1364 | 1362 | 1362 | C-H combination | CH ₃ |
| 1426 | 1426 | 1430 | | 1440 | 1424 | O-H stretch 1 st overtone | |
| | | | 1452 | | | O-H stretch 1 st overtone | H ₂ O, starch |
| 1474 | 1482 | 1478 | | | | O-H stretch 1 st overtone (intramolecular H-bond) | glucose |
| | | | | 1572 | | N-H stretch 1 st overtone | -CONH- |
| 1586 | 1586 | 1580 | | | | O-H stretch 1 st overtone (intermolecular H-bond) | Glucose, starch |
| 1672 | | | 1672 | | 1690 | =C-H | aromatic |
| 1722 | 1720 | 1722 | 1728 | 1724 | 1730 | C-H stretch 1 st overtone | CH ₂ |
| | | | | 1762 | | C-H stretch 1 st overtone | CH ₂ |
| 1784 | 1780 | 1774 | | | | C-H stretch/HOH deformation | cellulose |

Table 6 (Cont'd)

| Wood | NIR band (nm) | | | | | Assignment | Structure |
|------|---------------|---------------------|--------|------------|------|---|---------------------------|
| | Holocellulose | α -Cellulose | Lignin | Extractive | Ash | | |
| | 1826 | 1828 | | 1822 | 1830 | O-H stretch/C-O stretch 2 nd overtone | cellulose |
| 1920 | 1920 | 1920 | 1920 | 1920 | 1920 | O-H stretch/ O-H deformation 2 nd overtone | |
| | | | 2050 | 2064 | 1992 | | |
| 2084 | 2094 | 2092 | | | | Asym C-O-O stretch 3 rd overtone | cellulose, starch |
| | | | | | 2110 | N-H symmetric stretching | CONH ₂ , CONHR |
| 2132 | | | 2136 | 2134 | | =C-H stretch/ C=C stretch | HC=CH |
| | | | | | 2234 | | |
| 2268 | 2268 | 2268 | 2268 | 2270 | | O-H stretch/ C-O stretch | Cellulose, Lignin |
| | | | | | 2304 | | |
| 2330 | 2330 | 2338 | 2338 | 2342 | 2338 | C-H stretch/ C-H deformation | Cellulose, Lignin |
| | | | 2380 | | | O-H deformation 2 nd overtone | ROH |

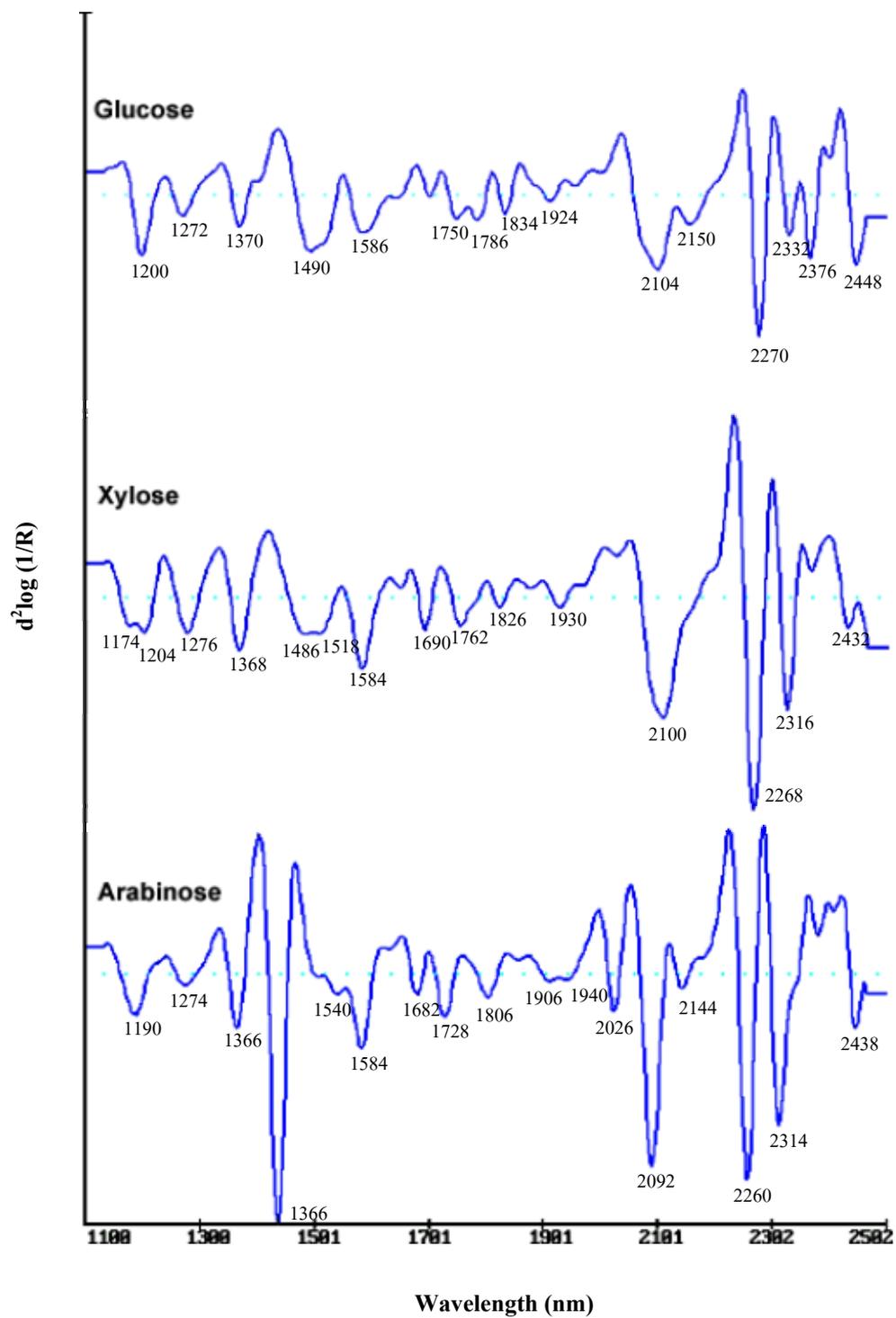


Figure 39 Second derivative spectra of glucose, xylose, arabinose standard.

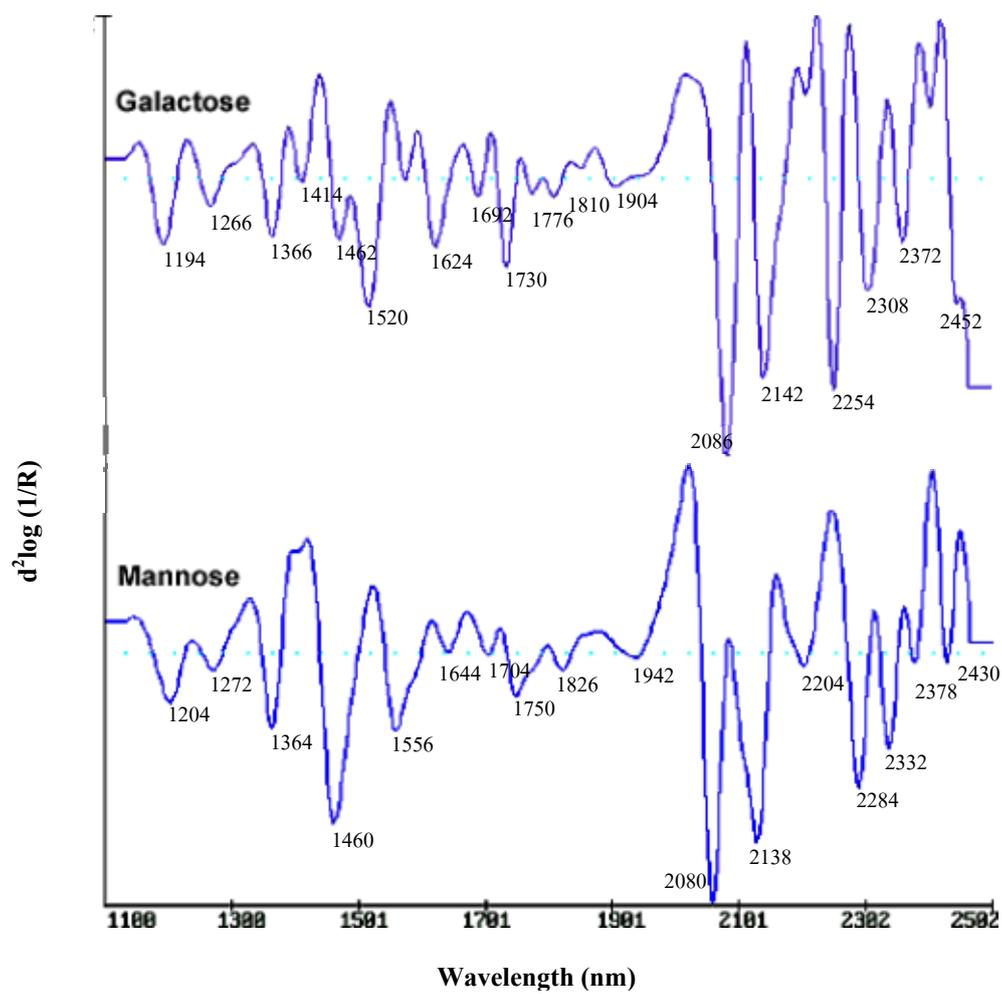


Figure 40 Second derivative spectra of galactose, mannose standard.

Table 7 Tentative assignment of NIR bands for monosaccharide standards (Burn,2001, Ozaki,2001,Osborn,1997, Shimleck,1997).

| NIR band (nm) | | | | | Assignment | Structure |
|---------------|--------|-----------|-----------|---------|---|-------------------|
| Glucose | Xylose | Arabinose | Galactose | Mannose | | |
| | 1174 | | | | C-H stretch 2 nd overtone | CH=CH |
| 1200 | 1204 | 1190 | 1194 | 1204 | C-H stretch 2 nd overtone | CH ₃ |
| 1272 | 1276 | 1274 | 1266 | 1272 | | |
| 1370 | 1368 | 1366 | 1366 | 1364 | C-H combination | CH ₃ |
| | | | 1414 | | 2 x C-H stretch + C-H deformation | CH ₂ |
| | | | 1462 | 1460 | | |
| 1490 | 1486 | | | | O-H stretch 1 st overtone (intramolecular H-bond) | cellulose, starch |
| | 1510 | | 1520 | | O-H stretch 1 st overtone | |
| | | 1540 | | 1556 | O-H stretch 1 st overtone | |
| 1586 | 1584 | 1584 | | | O-H stretch 1 st overtone (intermolecular H-bond) | CH ₂ |
| | | 1682 | 1624 | 1644 | | |
| | 1690 | | 1692 | | C-H stretch 1 st overtone | CH ₃ |

Table 7 (Cont'd)

| NIR band (nm) | | | | | Assignment | Structure |
|---------------|--------|-----------|-----------|---------|--|----------------------|
| Glucose | Xylose | Arabinose | Galactose | Mannose | | |
| | | | | 1704 | C-H stretch 1 st overtone | CH ₃ |
| | | 1728 | 1730 | | C-H stretch 1 st overtone | CH ₃ |
| 1750 | 1762 | | | 1750 | C-H stretch 1 st overtone | CH ₂ |
| 1786 | | | 1776 | | C-H stretch 1 st overtone | |
| 1834 | 1826 | 1806 | 1810 | 1826 | O-H stretch/C-O stretch 2 nd overtone | |
| | | 1906 | 1904 | | C=O stretch 2 nd overtone | |
| 1924 | 1930 | | | | O-H stretch/HOH deformation | |
| | | 1940 | | 1942 | O-H bend 2 nd overtone | H ₂ O |
| | | 2026 | | | | |
| 2104 | 2100 | 2092 | 2086 | 2080 | O-H stretch/HOH deformation | ROH, sucrose, starch |
| 2150 | | 2144 | 2142 | 2138 | C-H stretch/C=O stretch | |
| | | | | 2204 | | |

Table 7 (Cont'd)

| NIR band (nm) | | | | | Assignment | Structure |
|---------------|--------|-----------|-----------|---------|--|-----------|
| Glucose | Xylose | Arabinose | Galactose | Mannose | | |
| | | | 2254 | | O-H stretch/O-H deformation | |
| 2270 | 2268 | 2260 | | | O-H stretch/C-O stretch | |
| | | | | 2284 | C-H stretch/ CH ₂ deformation | |
| | 2316 | 2314 | 2308 | | C-H bending 2 nd overtone | |
| 2332 | | | | 2332 | C-H stretch/ C-H deformation | |
| 2376 | | | 2372 | 2378 | O-H deformation 2 nd overtone | ROH |
| 2448 | 2432 | 2438 | 2452 | 2430 | | |

4. The Data Analysis

Spectral data were analysed by using multiple linear regressions (MLR) and partial least square regression (PLSR) analysis method to develop a relationship between the value from conventional method and the value from NIR spectral data. MLR was performed by NSAS of Foss NIR Systems, Silver Spring, U.S.A. software. Unscramble software from CAMO AS, Norway was used for PLSR calculation.

Data sets comprising of the NIR spectra were split into two approximate subsets which used for modeling data sets; a calibration set and validation set, for establishing a predictive relationship between the spectra and chemical property of interest. The splitting was 42 samples for calibration set and then 28 samples of validation set.

The results of relationship between chemical component properties, range of content, average content and standard deviation were shown in Table 8 and Table 9.

Table 8 Relationships of chemical component property, range of content, average of standard derivation of 70 samples.

| Property | Group | Number | Range* | Average* | SD* |
|---------------------|--------------|---------------|---------------|-----------------|------------|
| Total pulp yield | Cal | 42 | 38.02-50.63 | 46.13 | 2.97 |
| | Val | 28 | 39.33-50.16 | 45.74 | 2.80 |
| Screen pulp yield | Cal | 42 | 37.94-50.43 | 46.04 | 2.98 |
| | Val | 28 | 39.04-50.10 | 45.64 | 2.83 |
| α -Cellulose | Cal | 42 | 39.14-48.93 | 44.04 | 2.39 |
| | Val | 28 | 40.11-46.65 | 43.20 | 1.69 |
| Holocellulose | Cal | 42 | 73.82-83.06 | 79.44 | 2.23 |
| | Val | 28 | 74.06-82.62 | 79.03 | 2.08 |
| Pentosan | Cal | 42 | 12.64-17.86 | 15.49 | 1.32 |
| | Val | 28 | 12.91-16.83 | 15.22 | 1.04 |
| Lignin | Cal | 42 | 25.42-32.12 | 28.54 | 1.32 |
| | Val | 28 | 26.59-29.43 | 28.15 | 0.75 |

* Unit (%W/W)

Table 8 (Cont'd)

| Property | Group | Number | Range* | Average* | SD* |
|--------------------|--------------|---------------|---------------|-----------------|------------|
| 1% NaOH solubility | Cal | 42 | 13.71-26.80 | 18.68 | 3.17 |
| | Val | 28 | 13.94-22.05 | 17.84 | 2.60 |
| Extractive | Cal | 42 | 0.80-4.31 | 2.51 | 0.96 |
| | Val | 28 | 1.04-3.96 | 2.33 | 0.91 |
| Ash | Cal | 42 | 0.61-1.63 | 0.87 | 0.22 |
| | Val | 28 | 0.63-1.04 | 0.79 | 0.12 |

* Unit (%W/W)

Table 9 Relation of sugar property, range of content, range of content and standard deviation of 70 samples.

| Property | Group | Number | Range* | Average* | SD* |
|-----------------|--------------|---------------|---------------|-----------------|------------|
| Glucose | Cal | 42 | 43.74-52.69 | 48.49 | 1.99 |
| | Val | 28 | 45.91-51.08 | 48.03 | 1.45 |
| Xylose | Cal | 42 | 8.81-16.57 | 13.09 | 2.37 |
| | Val | 28 | 9.16-15.89 | 12.73 | 2.18 |
| Arabinose | Cal | 42 | 0.38-1.01 | 0.63 | 0.17 |
| | Val | 28 | 0.40-0.91 | 0.59 | 0.14 |
| Galactose | Cal | 42 | 1.16-3.67 | 2.24 | 0.57 |
| | Val | 28 | 1.51-3.07 | 2.13 | 0.04 |
| Mannose | Cal | 42 | 0.28-1.55 | 0.70 | 0.25 |
| | Val | 28 | 0.43-1.03 | 0.63 | 0.13 |

* Unit (%W/W)

4.1 Total Pulp Yield

4.1.1 Multiple Linear Regression (MLR) Analysis of Spectral Data and Total Pulp Yield

The NSAS software was used for handling the NIR spectra. Many different multiple linear regression models were developed using the step up-search technique, which selected each wavelength that correlate to y-value. The numbers of best selected wavelengths were kept as much as possible which close to 4-5 numbers to avoid so called over fitting, which dramatically reduces the prediction accuracy of developed model. The result was not only increase the correlation coefficient (R), but also increase the value of standard error of prediction (SEP). In NSAS software for the maximum of MLR model, the number of selected wavelengths initially was set up at 4 wavelengths. The procedure of step-up search in MLR was applied to find the best wavelengths based on the comparison of correlation coefficients of X value of X matrix at for 701 of wavelengths with the selected constituent Y.

The best model for each property was pretreated with second derivative. Then was selected based on the values of correlation coefficient (R), standard error of calibration (SEC), standard error of prediction (SEP), and bias. In all cases, the derivative corrections of spectral data were used in order to extract important information from the pretreated absorbance data. The correlation at first selection of total pulp yield presents in Figure 41.

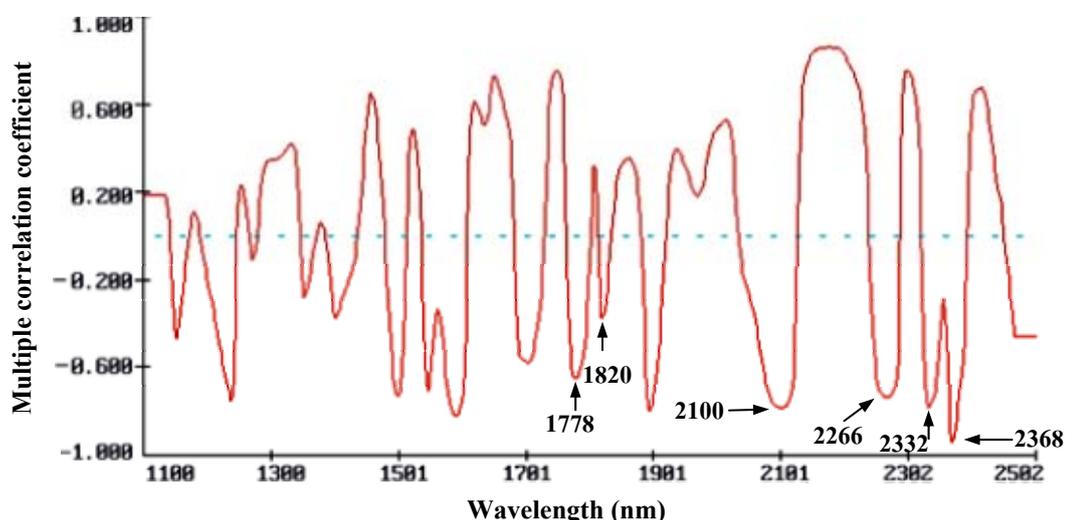


Figure 41 Second derivative multiple correlation coefficients of total pulp yield for select the first wavelength in set up calibration model.

In the correlation between wavelength (x-value) and total pulp yield (y-value), the wavelength at 2368 nm was highest correlation. Moreover, the wavelengths at 1778, 1820, 2100, 2266, and 2332 nm showed high correlations and had been reporting to cellulose band (Burn, 2001). The first step selects all of these wavelengths to form calibration model, and the second step use the computer for

selection by step-up search. The results were reported in appendix table A1. The percentage of total yield contents raised from the absorbance band at 1820 nm that was related to cellulose band combine with absorbance band at 2368, 1238, and 2152 nm, respectively presented the best statistic. The multiple correlation coefficients (R), SEC, SEP, and RPD value were 0.964, 0.833, 0.845, and 3.178, respectively. The calibration equation for total pulp yield analysis from NSAS program is;

$$\begin{aligned} \% \text{ Total pulp yield} = & 69.474 - 3288.676 d^2 \log(1/R_{1820}) - 3831.696 d^2 \log(1/R_{2368}) \\ & - 9329.276 d^2 \log(1/R_{1238}) + 882.995 d^2 \log(1/R_{2152}) \end{aligned} \quad (21)$$

This equation was used for predicted value that correlation with actual value. The correlation plots between the predicted value and actual value presented in Figure 42.

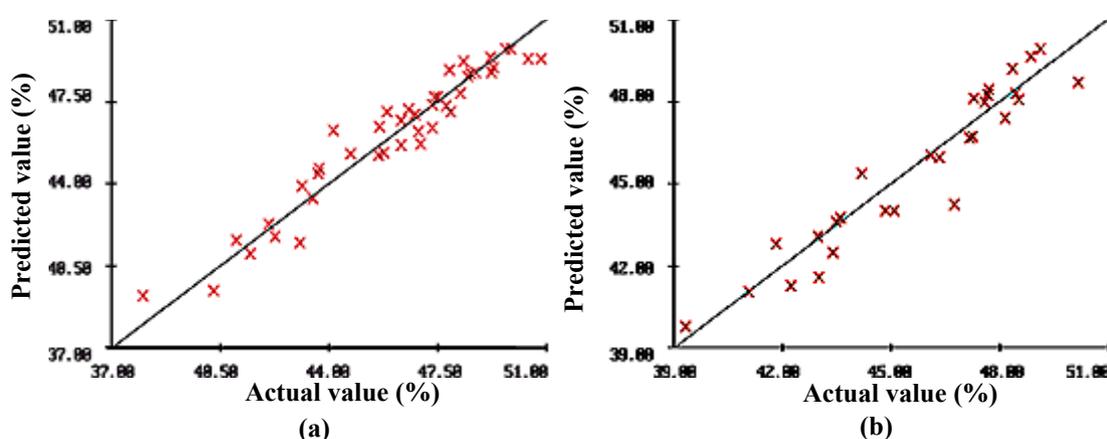


Figure 42 The correlation plots of total pulp yield content between predicted value and the actual value of calibration set (a) and validation set (b).

4.1.2 Partial Least Square Regression (PLSR) Analysis of Spectral Data and Total Pulp Yield

The PLSR method worked by reducing data dimensions to a number of factors. Using the calibration set only, predictive models were developed by PLSR method in which a complex multivariate relationship was established between the wood spectra and the values. A Number of pretreatment options were MSC, second derivative and the both pretreatment as a transformation to the absorbance spectra.

First step, the whole range of 1100-2500 nm spectra was pretreated with MSC, second derivative, and combinations to find the best pretreatment before made calibration of total pulp yield. The statistical results summarized in Table 10.

Table 10 The statistical summaries of total pulp yield content calibration model with different pretreatment methods.

| No. | Pretreatment | Factor | Calibration | | Prediction | | RPD |
|-----|------------------------|--------|-------------|-------|------------|--------|-------|
| | | | R | SEC* | SEP* | Bias | |
| 1 | Average spectra | 2 | 0.885 | 1.388 | 1.125 | 0.181 | 2.489 |
| 2 | MSC | 3 | 0.913 | 1.213 | 1.126 | 0.091 | 2.487 |
| 3 | Second derivatives | 2 | 0.939 | 1.021 | 0.971 | 0.135 | 2.884 |
| 4 | MSC+ second derivative | 1 | 0.918 | 1.183 | 1.084 | -0.078 | 2.583 |

* Unit (%W/W)

From the result, the second derivatives spectra was the best result for making the calibration, compared with offers pretreatment because the percentage value of SEC was lowest. Furthermore, selection of the possible wavelength that calculated by moving window statistic in math lab program. The result in plotting between log (SSR) and wavelength was shown in Figure 43.

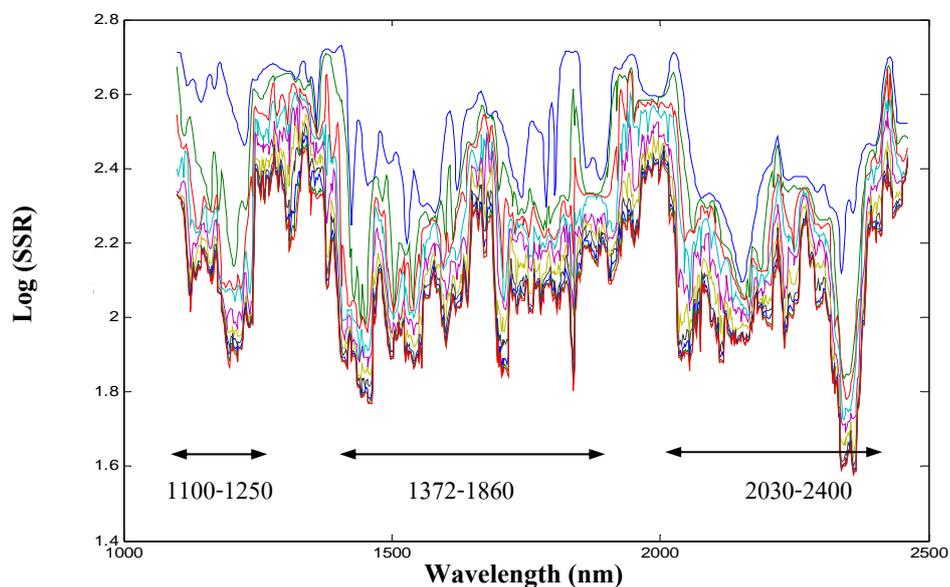


Figure 43 The residue lines obtained by MWPLSR for total pulp yield.

There were three regions which have low (SSR) value that correlated in the region of 1100-1250 nm, 1370-1860 nm and 2030-2400 nm, respectively. After that, these three regions were used to calculate the calibration. One region combined two regions and combined three regions were developed models of PLSR.

Table 11 The statistical summaries of total pulp yield content calibration model with vary three wavelength regions.

| No. | Region | Factor | Calibration | | Prediction | | RPD |
|-----|---------------------------------|--------|-------------|-------|------------|--------|-------|
| | | | R | SEC* | SEP* | Bias | |
| 1 | 1100-2500 | 2 | 0.939 | 1.021 | 0.971 | 0.135 | 2.884 |
| 2 | 1100-1250 | 3 | 0.944 | 0.986 | 0.991 | -0.071 | 2.825 |
| 3 | 1370-1860 | 2 | 0.940 | 1.020 | 1.065 | 0.172 | 2.629 |
| 4 | 2030-2400 | 3 | 0.954 | 0.897 | 1.058 | -0.023 | 2.647 |
| 5 | 1100-1250 +1370-1860 | 2 | 0.940 | 1.019 | 0.945 | 0.064 | 2.963 |
| 6 | 1100-1250 +2030-2390 | 3 | 0.947 | 0.954 | 0.994 | 0.021 | 2.817 |
| 7 | 1372-1860 + 2030-2400 | 2 | 0.938 | 1.032 | 0.983 | 0.085 | 2.848 |
| 8** | 1100-1250+1372-1860 + 2030-2390 | 2 | 0.935 | 1.057 | 0.949 | 0.066 | 2.950 |

* Unit (%W/W), ** Selected Calibration

The best results among developed models of PLSR for total pulp yield of *Eucalyptus* wood powder was used in 3 regions, 1100-1250, 1372-1860, and 2030-2390 nm, together. The correlation coefficient and standard error of calibration set were 0.935 and 1.057, the standard error of prediction and RPD were 0.949, and 2.950, respectively. The summaries of statistic results of developed equation showed in Table 12. The best calibration was obtained as equation 22.

$$\% \text{ Total pulp yield} = b_0 + \sum_{i=1100}^{1250} b_i X_i + \sum_{i=1372}^{1860} b_i X_i + \sum_{i=2030}^{2390} b_i X_i \quad (22)$$

Where b_0 is y-intercept of regression model, b is regression coefficient at wavelength i , X_i is $d^2 \log(1/R_i)$, i is wavelength

Regression coefficient plots were primarily used to check the importance of the difference absorbance (x-variable) for predicting the reference value (Y-value). Large absolute values indicate high important (significant) and low value (close to 0) indicate an unimportance variable. The coefficient value indicated the average increasing in y-value when the corresponding X-variable was increasing by one unit, which kept the other entire variables constant (Camo AS, 1996). The regression coefficient of this equation reported in Figure 44.

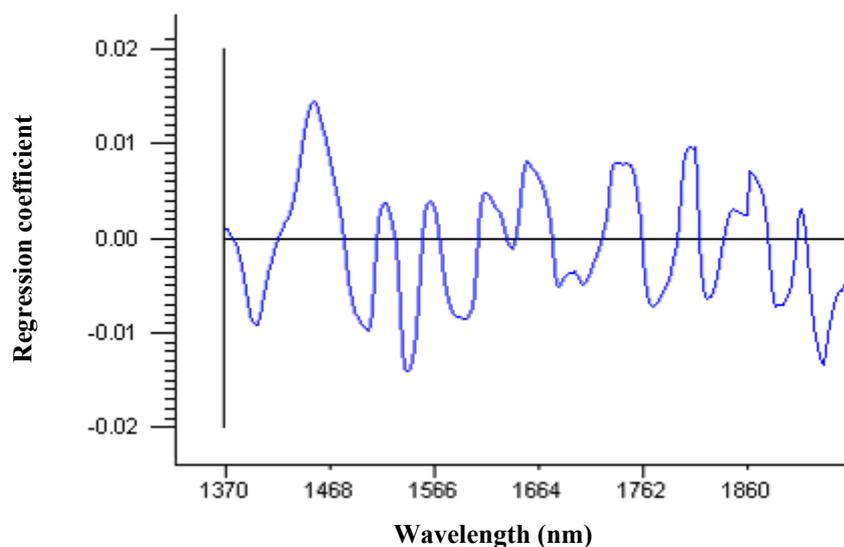


Figure 44 The regression coefficients from equation 22 of total pulp yield.

The correlation plots between the predicted value and actual value of calibration set and prediction set were illustrated in Figure 45.

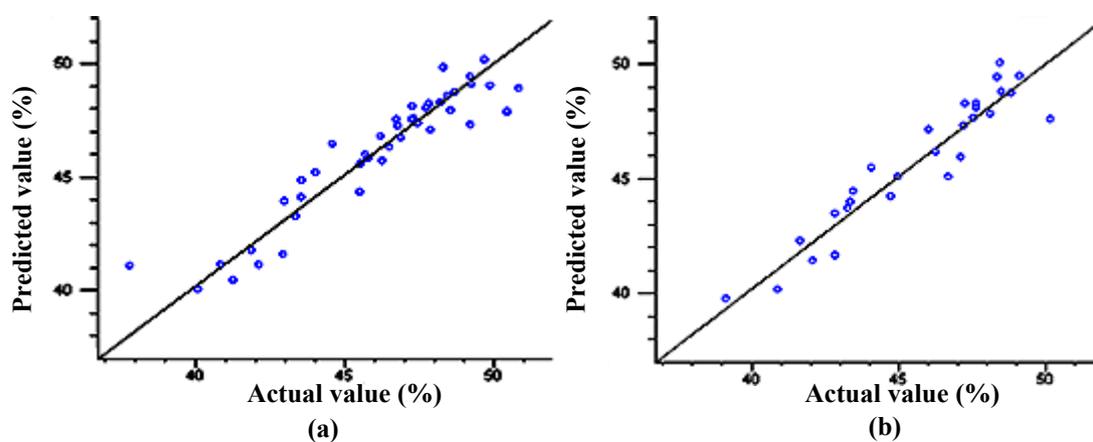


Figure 45 The correlation plots of total pulp yield content between predicted value and the actual value of calibration set (a) and validation set (b).

4.2 Screen Pulp Yield

4.2.1 Multiple Linear Regression (MLR) Analysis of Spectral Data and Screen Pulp Yield

The best model for each constituent was pretreated with second derivative. In all cases, the derivative corrections of spectral data were used in order to extract important information from the pretreated absorbance data. The correlations of the first wavelength selection were presented in Figure 46.

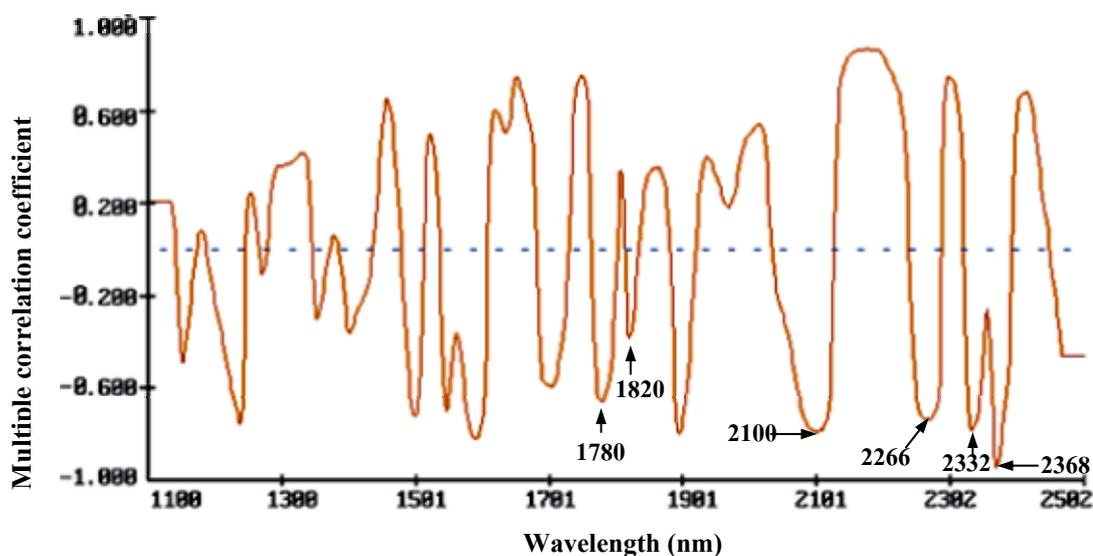


Figure 46 The multiple correlation coefficients of screen pulp yield that pretreated with second derivative.

Most of the correlations of the screen pulp yield value were the same as the total pulp yield value. Six wavelengths; 1780, 1820, 2100, 2266, 2332, 2368 nm, were selected in the same way as total pulp yield that had been reporting from Burn (2001). The wavelength at 2368 nm was the highest correlation. The band at 1780, 2100, 2266, 2332, 2368 nm had been reporting to cellulose band as the same as Burn (2001). The result was shown in appendix table A2.

The result also showed in the same of the total pulp yield. The first absorbance at 1820 nm follows with absorbance at 2368, 1238, and 2152, respectively. The equation performed the best statistic result. The multiple correlation coefficients (R), SEC, SEP, and RPD were 0.967, 0.804, 0.843, and 3.357, respectively. The correlation plot of the validation set and the calibration set presented in Figure 47. The calibration for screen pulp yield analysis from NSAS program was obtained as

$$\% \text{ Screen pulp yield} = 69.302 + 3042.355d^2\log(1/R_{1820}) - 3894.900d^2\log(1/R_{2368}) - 9211.274 d^2\log(1/R_{1238}) - 861.279d^2\log(1/R_{2152}) \quad (23)$$

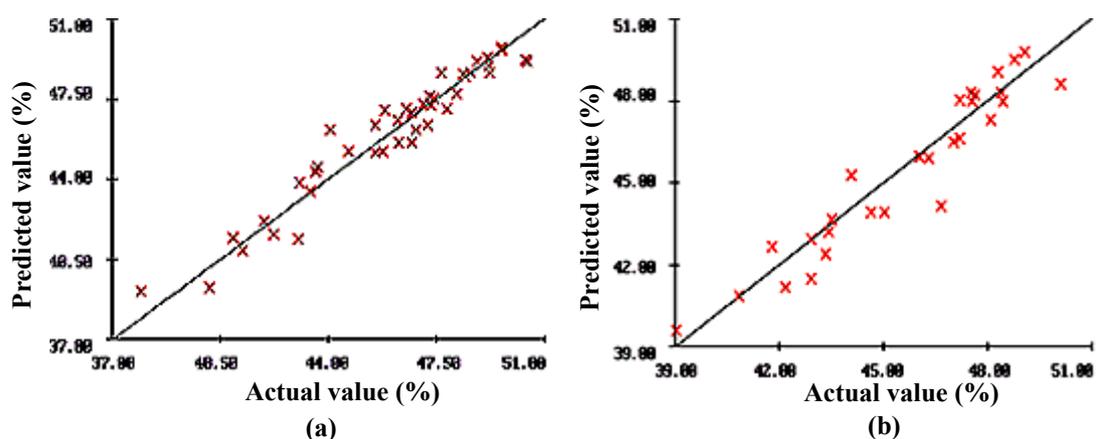


Figure 47 The correlation plots of screen pulp yield between predicted value and actual value of calibration set (a) and validation set (b)

4.2.2 Partial Least Square Regression (PLSR) Analysis of Spectral Data and Screen Pulp Yield

This analysis was compared between using the pretreatment data before making calibration for the screen pulp yield value. The whole wavelengths in the range of 1100-2500 nm were calculated and validated as test set method. The statistic summary results were presented in Table 12.

Table 12 The summaries of statistic in screen pulp yield content calibration model with different pretreatment method.

| No. | Pretreatment | Factor | Calibration | | Prediction | | RPD |
|-----|------------------------|--------|-------------|-------|------------|--------|-------|
| | | | R | SEC* | SEP* | Bias | |
| 1 | Average spectra | 2 | 0.889 | 1.361 | 1.125 | 0.206 | 2.516 |
| 2 | MSC | 2 | 0.906 | 1.26 | 1.197 | 0.244 | 2.364 |
| 3 | Second derivatives | 2 | 0.939 | 1.026 | 0.969 | -0.106 | 2.921 |
| 4 | MSC+ second derivative | 1 | 0.869 | 1.321 | 1.047 | -0.043 | 2.703 |

* Unit (%W/W)

The result presented to the second derivatives spectra was provided the best calibration model. The SEC, SEP and bias were the lowest value. The R and RPD were highest. Then these spectra were transferred to moving window program. The result between log (SSR) and wavelength was shown in Figure 48.

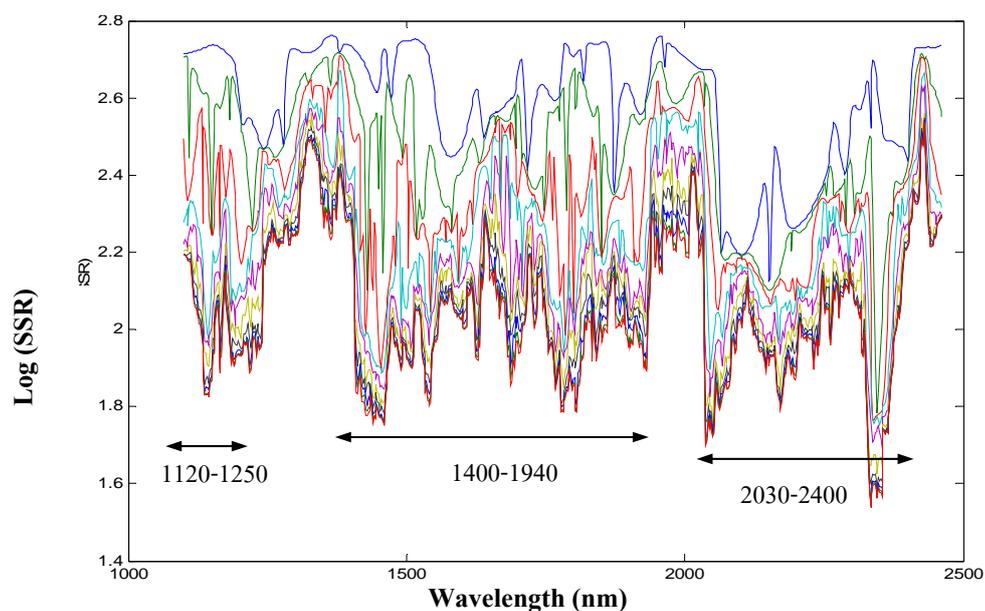


Figure 48 The residue lines obtained by MWPLSR for screen pulp yield.

The spectra were pretreated with second derivative, then export to Moving Window Selection program. The residual line presented in three regions that have low log (SSR) value. The region in the range of 2030-2400 nm presented the lowest log (SSR). The region in the range of 1100-1250 nm indicated low log (SSR) too, the combination of two and three regions estimated the screen pulp yield was calculated. The statistical result presented in Table 13.

Table 13 The summaries of statistic in screen pulp yield content calibration model with vary three wavelength regions.

| No. | Region | Factor | Calibration | | Prediction | | RPD |
|-----|-----------------------------------|--------|-------------|-------|------------|--------|-------|
| | | | R | SEC* | SEP* | Bias | |
| 1 | 1100-2500 | 2 | 0.939 | 1.026 | 0.969 | -0.106 | 2.921 |
| 2 | 1100-1250 | 4 | 0.953 | 0.903 | 0.984 | -0.096 | 2.876 |
| 3 | 1400-1940 | 2 | 0.935 | 1.054 | 1.021 | 0.046 | 2.772 |
| 4 | 2030-2400 | 2 | 0.942 | 1.001 | 1.097 | -0.058 | 2.580 |
| 5 | 1100-1250 + 1400-1940 | 2 | 0.936 | 1.045 | 0.953 | 0.008 | 2.970 |
| 6** | 1100-1250+2030-2400 | 3 | 0.963 | 0.799 | 0.938 | -0.078 | 3.017 |
| 7 | 1400-1940+2030-2400 | 2 | 0.944 | 0.979 | 1.064 | 0.107 | 2.660 |
| 8 | 1120-1250+1400-1940 +2030-2400 | 2 | 0.940 | 1.016 | 0.954 | 0.042 | 2.966 |

* Unit (%W/W), ** Selected Calibration

From table 13, the best calibration was consisted of the absorbance in range 1100-1250 nm and the region of 2030-2400 nm combination. The region of 1400-1940 nm contains two water bands at 1424 nm and 1918 nm which both band were not corresponded to build the calibration for prediction the screen yield. The correlation coefficient of calibration set is 0.963 and standard error of prediction is 0.938, as equation 24.

$$\% \text{ Screen pulp yield} = b_0 + \sum_{i=1100}^{1250} b_i x_i + \sum_{i=2030}^{2400} b_i x_i \quad (24)$$

Where b_0 is y-intercept of regression model, b is regression coefficient at wavelength i , X_i is $d^2 \log(1/R_i)$, i is wavelength. There were several values of the regression coefficients for equation 24, which was shown as Figure 49.

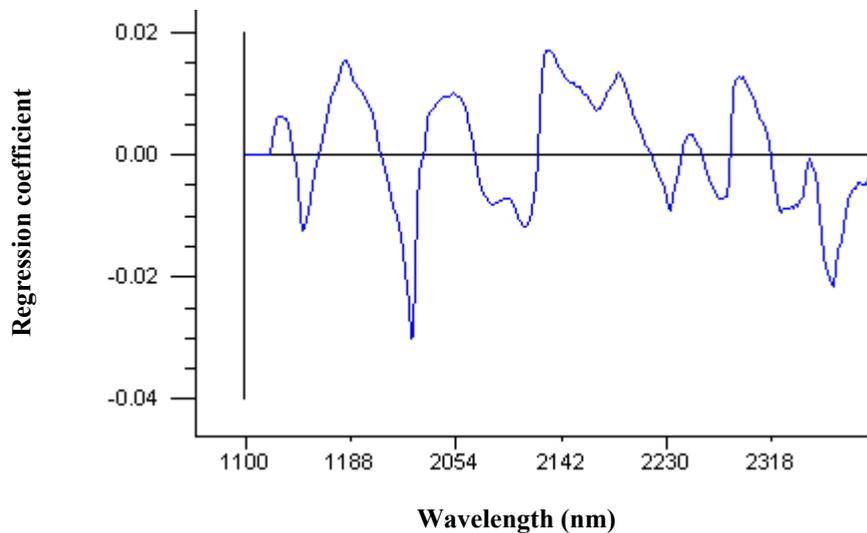


Figure 49 The regression coefficients from equation 24 with using wavelength in the region 1100-1250 nm and 2030-2400 nm

The correlation plots between the predicted value and actual value of the calibration set and the validation set were shown in Figure 50.

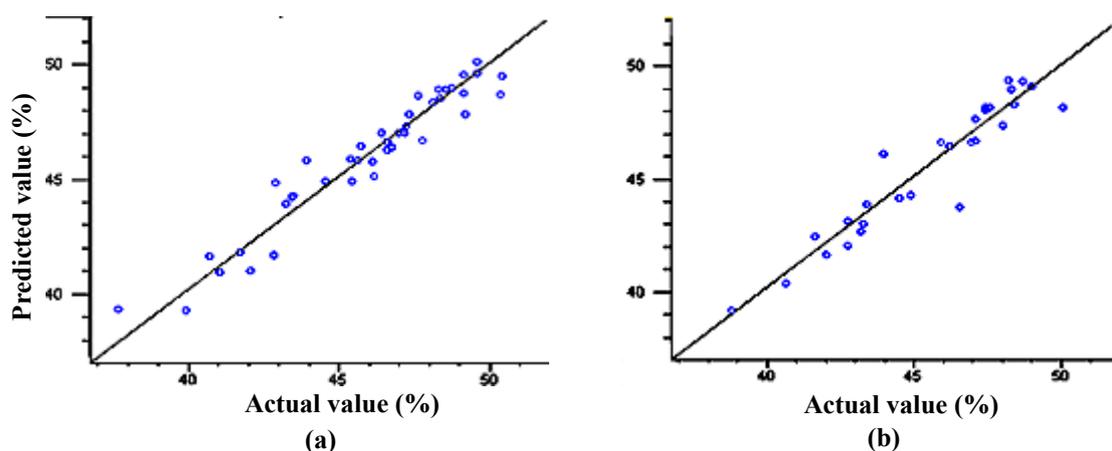


Figure 50 The correlation plots of screen pulp yield between predicted value and the actual value of calibration set (a) and validation set (b)

4.3 Holocellulose

4.3.1 Multiple Linear Regression (MLR) Analysis of Spectral Data and Holocellulose

The correlation of absorbance in wavelength range at 1100-2500 nm and holocellulose content in each wavelength was pretreated with second derivative to extract important information. The multiple correlations were presented in Figure 51.

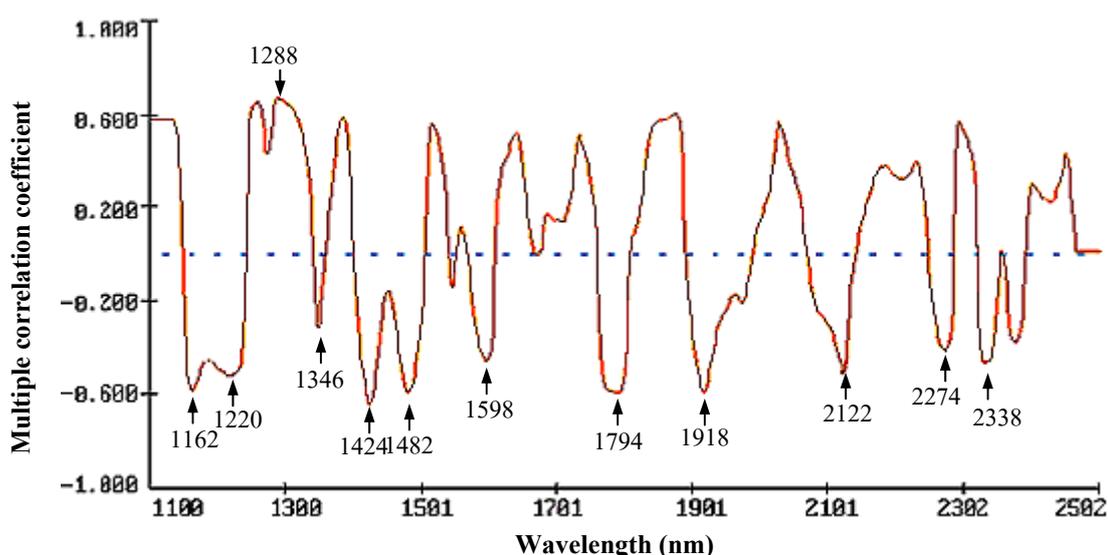


Figure 51 The multiple correlation coefficients of holocellulose that pretreated with second derivative.

The automatic selection selected absorbance at 1288 nm. The first selection wavelength method was similar to α -cellulose. The tentative assignment

bands were related to holocellulose bands. Some band was not appearing such as the band at 1820 nm. However, there were twelve bands that related to the second derivative spectrum of holocellulose standard band; 1162, 1220, 1288, 1346, 1424, 1482, 1598, 1794, 1918, 2122, 2274, and 2338 nm. The results showed in the appendix table A3. Three absorbance bands that provided the best calibration model for the holocellulose contents estimation. The first absorbance band at 2338 nm represented to C-H bond stretching and C-H bending combination of cellulose band. The others were the absorbance at 1474, 2176, 2454 nm, that gave the best statistic results. The best result was selected from the calibration model which was presented the lowest SEP. The multiple correlation coefficients (R), SEC, SEP, and RPD were 0.693, 1.690, 1.650, and 1.261, respectively. The correlation plot for validate the calibration presents in Figure 52. The best calibration was obtained as equation 25.

$$\begin{aligned} \% \text{ Holocellulose} = & 52.541 + 1655.485 d^2 \log(1/R_{2338}) - 2183.908 d^2 \log(1/R_{1474}) \\ & + 3017.759 d^2 \log(1/R_{2176}) + 1368.922 d^2 \log(1/R_{2454}) \end{aligned} \quad (25)$$

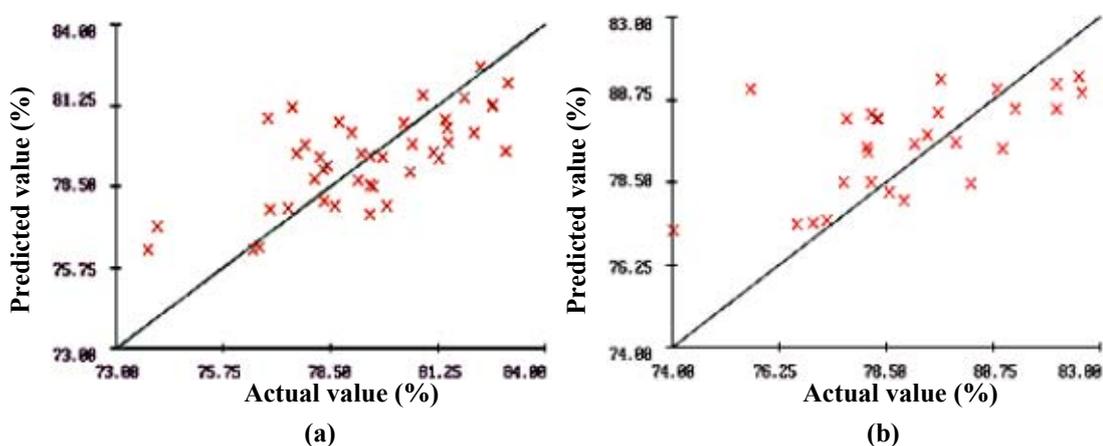


Figure 52 The correlation plots of holocellulose between predicted value and the actual value of calibration set (a) and validation set (b).

4.3.2 Partial Least Square Regression (PLSR) Analysis of Spectral Data and Holocellulose

This analysis compared between different pretreatment data before making calibration to estimate the holocellulose content. The whole wavelength in the region of 1100-2500 nm was used to calculate and validate as test set method. The statistical summaries of the result showed in Table 14.

Table 14 The statistical summaries of holocellulose calibration model with different pretreatment methods.

| No. | Pretreatment | Factor | Calibration | | Prediction | | RPD |
|-----|------------------------|--------|-------------|-------|------------|-------|-------|
| | | | R | SEC* | SEP* | Bias | |
| 1 | Average spectra | 3 | 0.715 | 1.561 | 1.808 | 0.547 | 1.151 |
| 2 | MSC | 2 | 0.693 | 1.590 | 1.737 | 0.514 | 1.198 |
| 3 | Second derivatives | 1 | 0.665 | 1.648 | 1.693 | 0.344 | 1.229 |
| 4 | MSC+ second derivative | 1 | 0.655 | 1.666 | 1.808 | 0.480 | 1.151 |

* Unit (%W/W)

The second derivative spectra were shown the lowest SEP value, and then transferred these spectra to moving window program. The result between log (SSR) and wavelength showed in Figure 53.

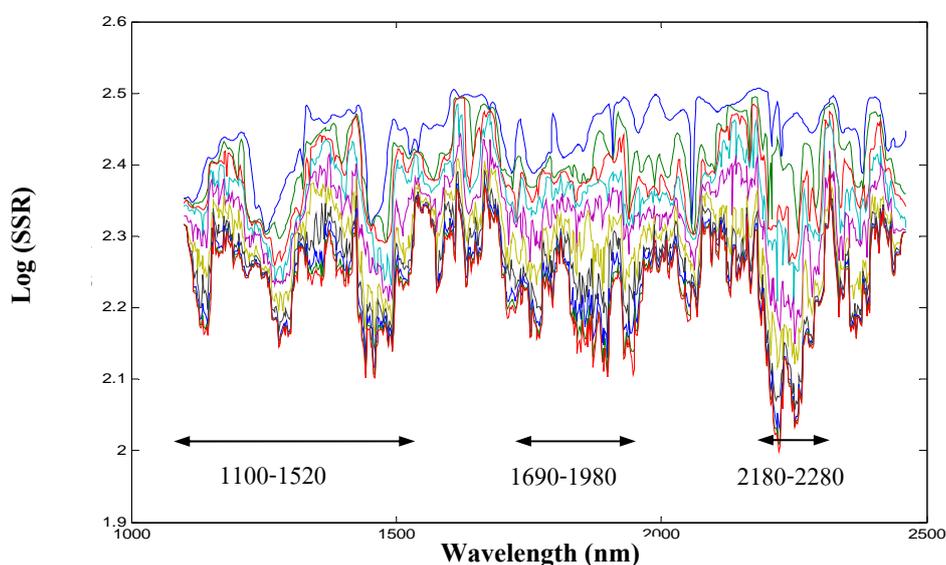


Figure 53 The residual lines obtained by MWPLSR for holocellulose.

The regressions were in the range of 1100-1520 nm, 1690-1980 nm and 2180-2280 nm. The combination of the absorbance of two and three regions was calculated for the holocellulose-content estimation was calculated. The statistical results showed in Table 15.

Table 15 The statistical summaries of holocellulose calibration model with vary three wavelength regions.

| No. | Region | Factor | Calibration | | Prediction | | RPD |
|-----|------------------------------------|--------|-------------|-------|------------|-------|-------|
| | | | R | SEC* | SEP* | Bias | |
| 1 | 1100-2500 | 1 | 0.665 | 1.648 | 1.693 | 0.344 | 1.229 |
| 2 | 1100-1500 | 2 | 0.673 | 1.652 | 1.858 | 0.392 | 1.120 |
| 3 | 1690-1980 | 2 | 0.613 | 1.764 | 1.657 | 0.220 | 1.256 |
| 4** | 2180-2300 | 7 | 0.779 | 1.400 | 1.496 | 0.245 | 1.391 |
| 5 | 1100-1500+1690-1980 | 1 | 0.670 | 1.657 | 1.768 | 0.402 | 1.177 |
| 6 | 1100-1500+2180-2300 | 1 | 0.668 | 1.662 | 1.827 | 0.363 | 1.139 |
| 7 | 1100-1500 + 1690-1980+2180-2300 | 1 | 0.667 | 1.664 | 1.753 | 0.380 | 1.187 |

* Unit (%W/W) ,** selected Calibration

The absorbance at the three regions: 1100-1500, 1690-1980, and 2180-2300 nm was developed the calibration model. The region of 2180-2300 nm gave the best calibration. The statistic result; the correlation coefficient, SEC of calibration set, the SEP, RPD of validation set were 0.779, 1.400, 1.496, and 1.246, respectively. The best calibration equation was obtained as equation 26.

$$\% \text{ Holocellulose} = b_0 + \sum_{i=2180}^{2300} b_i x_i \quad (26)$$

Where b_0 is y-intercept of regression model, b is regression coefficient at wavelength i , X_i is $d^2 \log(1/R_i)$, i is wavelength

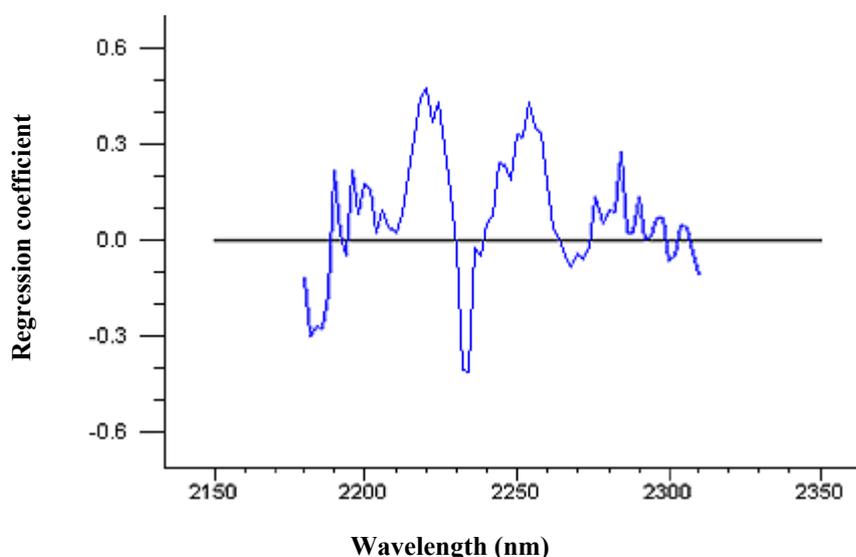


Figure 54 The regression coefficients from equation 26 with use in region 2180-2280 nm.

The regression coefficients of equation 26 were shown in Figure 54. The correlation plots between the predicted value and actual value was shown in Figure 55.

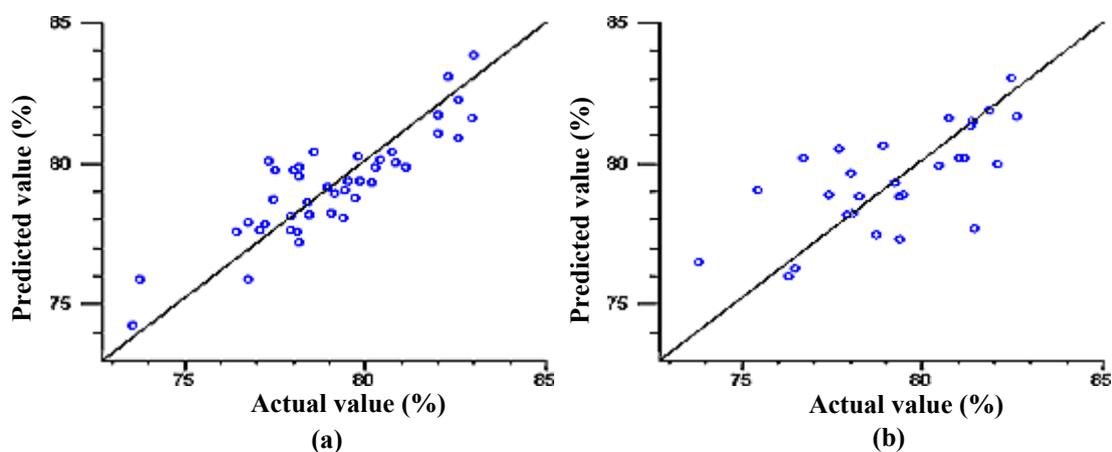


Figure 55 The correlation plots of holocellulose between predicted value and actual value of calibration set (a) and validation set (b).

4.4 α -Cellulose

4.4.1 Multiple Linear Regression (MLR) Analysis of Spectral Data and α -Cellulose

The correlation of absorbance at in range at 1100-2500 nm and α -cellulose content in each wavelength was pretreated with second derivative to extract important information. The multiple correlations were shown in Figure 56.

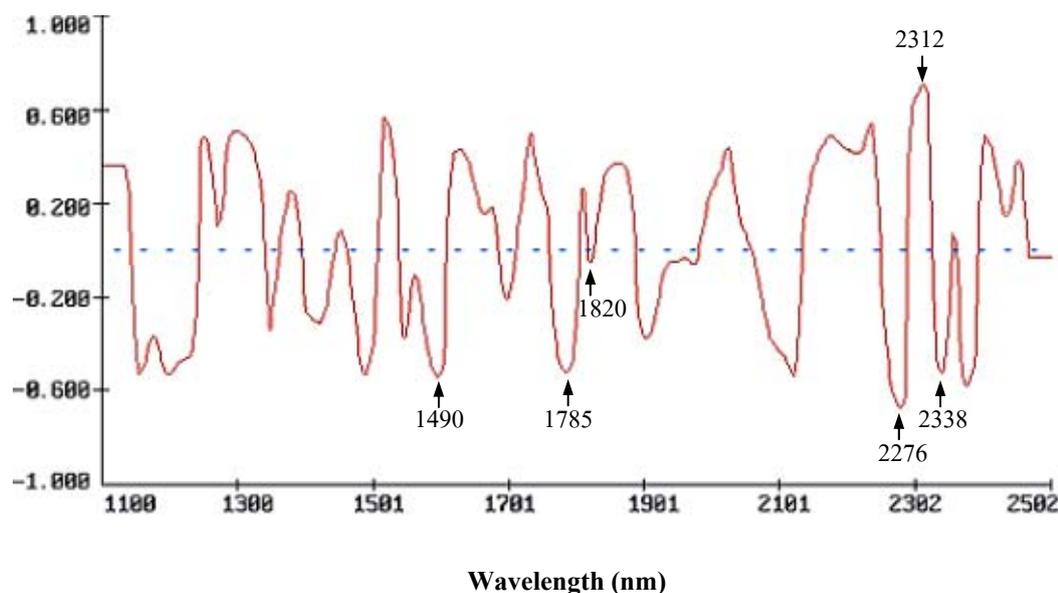


Figure 56 The multiple correlation coefficients of α -cellulose that pretreated with second derivative.

The automatic selection selected the first absorbance at 2312 nm that related to C-H 2nd overtone (Burn, 2001). In addition, the other band that related to the second derivative spectrum of α -cellulose standard was selected. The summaries of statistical result showed in appendix table A4. The best result used the three regions to estimate the α -cellulose content. The first absorbance was 2276 nm that relate to cellulose peak, next was the 1396 nm related to C-H combination. However, absorbance at 1688 related to aromatic compound, which was represented lignin content, which influent to celluloses content.

The best statistic results reported the lowest SEP. The multiple correlation coefficients (R), SEC, SEP, were 0.869, 1.220, and 0.677, respectively. The RPD was 2.496. The correlation plot for validate the calibration presented in Figure 57. The calibration for α -cellulose content analysis from NSAS program was shown as equation 27.

$$\begin{aligned} \% \alpha\text{-cellulose} = & 27.920 - 1120.817 d^2 \log(1/R_{2276}) + 4330.397 d^2 \log(1/R_{1688}) \\ & - 730.669 d^2 \log(1/R_{1396}) \end{aligned} \quad (27)$$

The correlation plots between the predicted value and actual value of the calibration set and the validation set were shown in Figure 57.

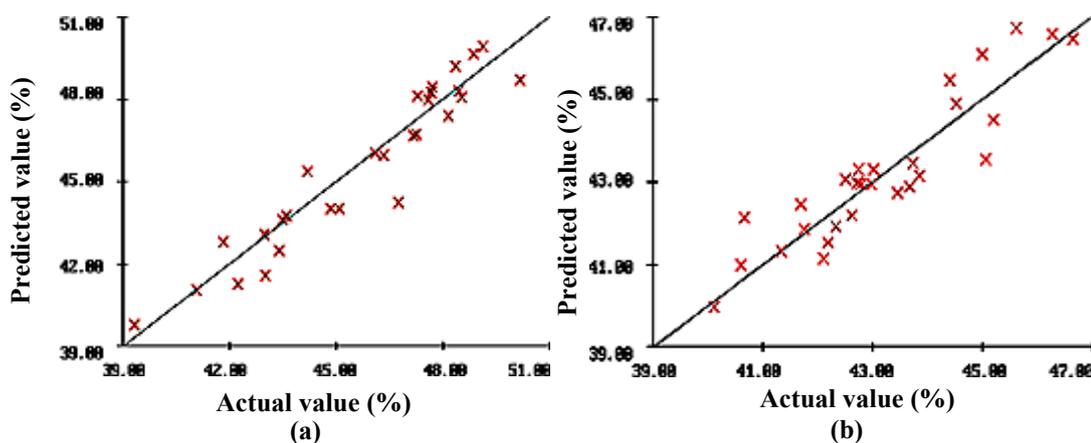


Figure 57 The correlation plots of α -cellulose between predicted value and the actual value of calibration set (a) and validation set (b).

4.4.2 Partial Least Square Regression (PLSR) Analysis of Spectral Data and α -Cellulose

This analysis compared between using the pretreatment before making calibration for α -cellulose content. The whole wavelength in the range of 1100-2500 nm was calculated and validated as test set method. The statistic summaries showed in Table 16.

Table 16 The statistical summaries of α -cellulose calibration model that vary the pretreatment.

| No. | Pretreatment | Factor | Calibration | | Prediction | | RPD |
|-----|------------------------|--------|-------------|-------|------------|--------|-------|
| | | | R | SEC* | SEP* | Bias | |
| 1 | Average spectra | 5 | 0.729 | 1.710 | 0.988 | 0.268 | 1.711 |
| 2 | MSC | 5 | 0.805 | 1.479 | 0.905 | 0.399 | 1.867 |
| 3 | Second derivatives | 5 | 0.913 | 0.962 | 0.772 | -0.103 | 2.189 |
| 4 | MSC+ second derivative | 5 | 0.909 | 0.985 | 0.819 | -0.084 | 2.063 |

* Unit (%W/W)

The result showed that, the spectra were pretreated with the second derivatives which presented the best result for making the calibration. The SEC and bias were lowest values, and then transferred these spectra to moving window program. The result between log (SSR) and wavelength was shown in Figure 58.

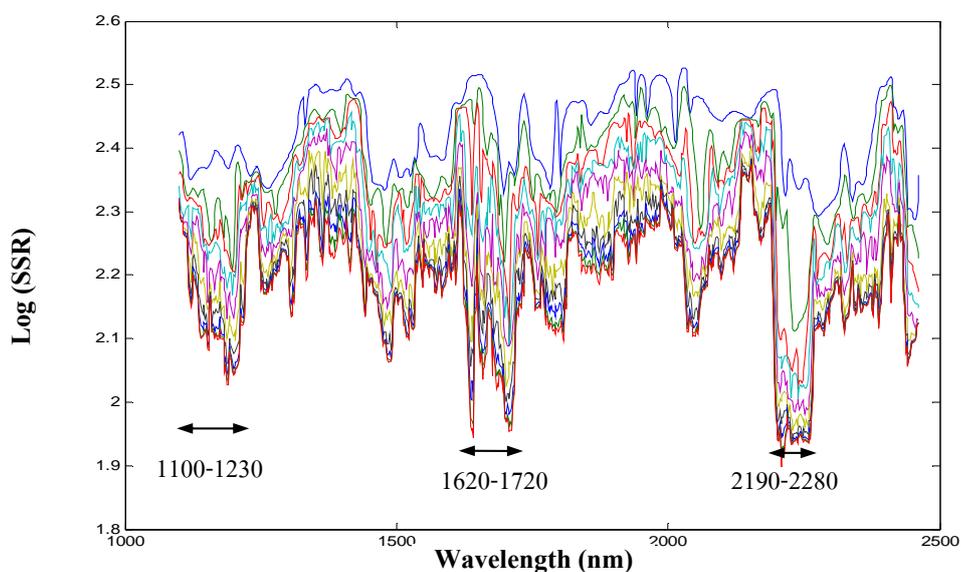


Figure 58 The residue lines obtained by MWPLSR for α -cellulose.

The three regions showed low log (SSR) value. The combination of the absorbance in the region of 1100-1230, 1620-1720, and 2190-2280 nm was estimated α -cellulose. The statistical results reported in Table 17.

Table 17 The statistical summaries of in α -cellulose calibration model that vary wavelength regions.

| No. | Region | Factor | Calibration | | Prediction | | RPD |
|-----|-----------------------------------|--------|-------------|-------|------------|--------|-------|
| | | | R | SEC* | SEP* | Bias | |
| 1 | 1100-1230 | 9 | 0.805 | 1.401 | 0.864 | -0.013 | 1.956 |
| 2 | 1620-1720 | 4 | 0.861 | 1.204 | 0.827 | 0.038 | 2.044 |
| 3 | 2190-2280 | 3 | 0.823 | 1.309 | 0.924 | 0.088 | 1.829 |
| 4 | 1100-1230+1620-1720 | 5 | 0.872 | 1.157 | 0.838 | 0.088 | 2.017 |
| 5 | 1100-1230+2190-2280 | 4 | 0.865 | 1.186 | 0.778 | 0.048 | 2.172 |
| 6 | 1620-1720+2190-2280 | 2 | 0.834 | 1.305 | 0.931 | 0.128 | 1.815 |
| 7 | 1100-1230+1620-1720 +2190-2280 | 3 | 0.872 | 1.157 | 0.838 | 0.088 | 2.017 |

* Unit (%W/W)

From table 17, the result was not better than the result from using the whole absorbance wavelength for building calibration. The best equation was calculated from whole absorbance. The correlation coefficient of calibration set, SEC, SEP, and RPD, were 0.913, 0.962, 0.772, and 2.189 respectively. The best calibration was obtained as equation 28.

$$\% \alpha - Cellulose = b_0 + \sum_{i=1100}^{2500} b_i x_i \quad (28)$$

Where b_0 is y-intercept of regression model, b is regression coefficient at wavelength i , X_i is $d^2 \log(1/R_i)$, i is wavelength

There were several values of the regression coefficients for equation 28 that were performed as Figure 59.

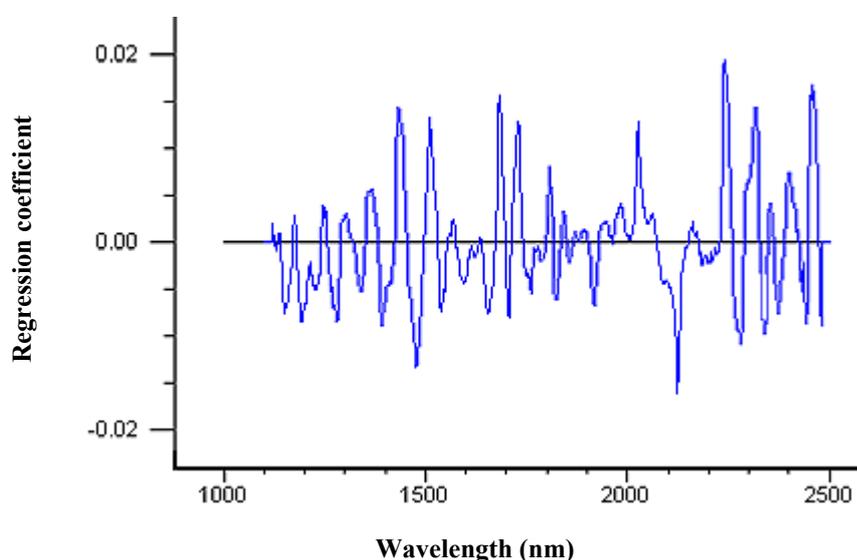


Figure 59 The correlation coefficients from equation 28 with using wavelengths in region of 1100-2500 nm.

The correlation plots between the predicted value and actual values of the calibration set and the validation set were illustrated in Figure 60.

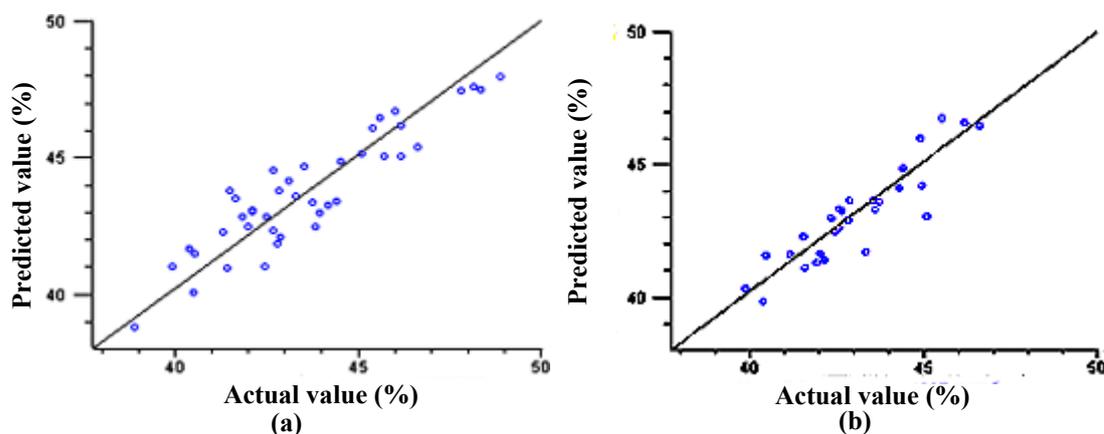


Figure 60 The correlation plots of α -cellulose between predicted value and the actual value of calibration set (a) and validation set (b).

4.5 Pentosan

4.5.1 Multiple Linear Regression (MLR) Analysis of Spectral Data and Pentosan

The absorbances in the region of 1100-2500 nm were pretreated with second derivative to extract important information. The multiple correlations were shown in Figure 61.

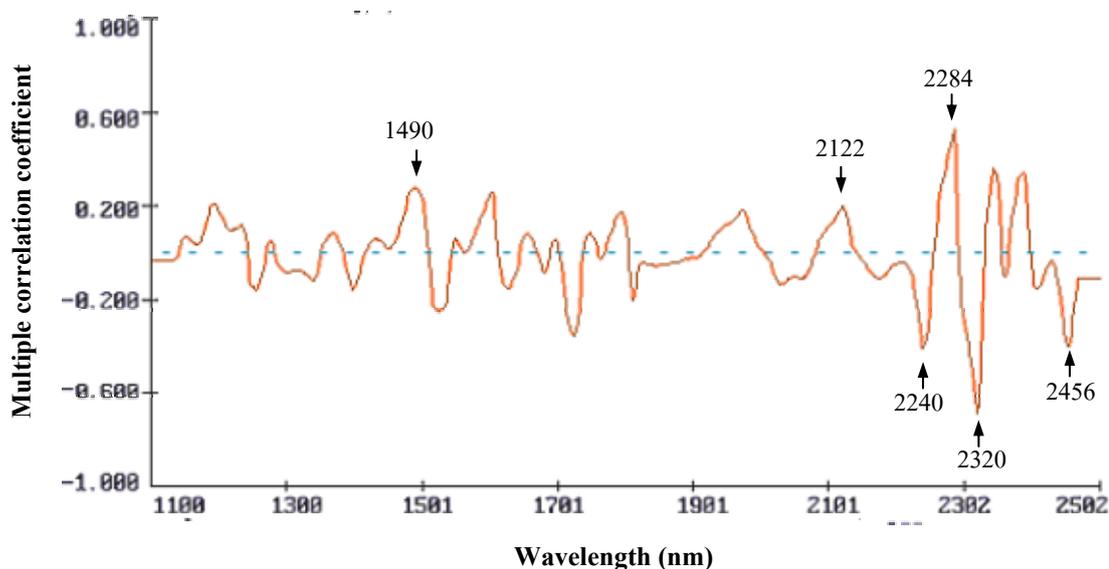


Figure 61 The multiple correlation coefficients of pentosan that pretreated with second derivative.

The multiple correlation bands that corresponded with the second derivative spectrum of xylose standard were selected for the first parameter in the calibration model. All of these bands were 1490, 2122, 2240, 2284, 2320, and 2456 nm. The best calibration arisen from absorbance at 2240 nm that related to xylan band (Michell, 1996). The band at 2320 nm related to the combination band between C-H stretching and CH₂ deformation and this band was dominant band of xylose standard. The band was 1604 nm didn't have the assignment. The correlation coefficient of calibration set, SEC, SEP, and RPD, were 0.772, 0.836, 0.761, and 1.367 respectively. The summary statistical results of developing networks were shown in appendix table 6. The best calibration was obtained as equation 29.

$$\begin{aligned} \% \text{ Pentosan} = & 6.916 - 176.219 d^2 \log(1/R_{2240}) - 1876.903 d^2 \log(1/R_{2320}) \\ & + 1644.519 d^2 \log(1/R_{1604}) \end{aligned} \quad (29)$$

The correlation plots shows between the predicted value and actual value, presents in Figure 62.

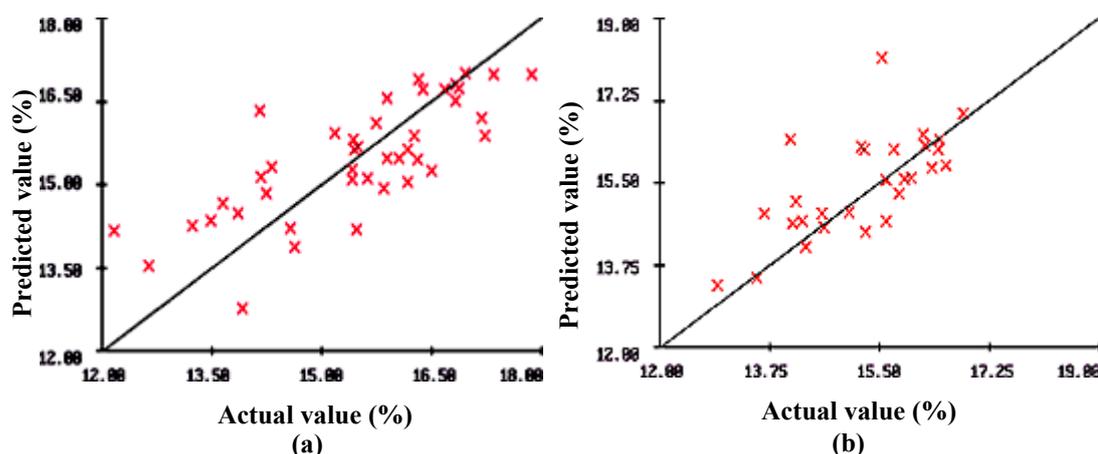


Figure 62 The correlation plots of pentosan between predicted value and the actual value of calibration set (a) and validation set (b).

4.5.2 Partial Least Square Regression (PLSR) Analysis of Spectral Data and Pentosan

This analysis compared between different pretreatment before calibration performing to estimate the pentosan content. The whole wavelength range in the region of 1100–2500 nm was used to calculate and validate as test set method. The statistical summaries showed in Table 18.

Table 18 The statistical summaries of pentosan calibration model with different pretreatment methods in vary method

| No. | Pretreatment | Factor | Calibration | | Prediction | | RPD |
|-----|------------------------|--------|-------------|-------|------------|-------|-------|
| | | | R | SEC* | SEP* | Bias | |
| 1 | Average spectra | 4 | 0.479 | 1.156 | 0.841 | 0.288 | 1.237 |
| 2 | MSC | 3 | 0.558 | 1.093 | 0.865 | 0.265 | 1.202 |
| 3 | Second derivatives | 3 | 0.782 | 0.820 | 0.795 | 0.262 | 1.308 |
| 4 | MSC+ second derivative | 1 | 0.596 | 1.057 | 0.795 | 0.219 | 1.308 |

* Unit (%W/W)

The SEC value from second derivative pretreatment calibration showed equal value from MSC and second derivative pretreatment calibration. So, the selection in pretreatment was only selected pretreatment with second derivative. After that Moving window selection was used for the possible absorbance region to make calibration. The resulting between log (SSR) and wavelength was presented in Figure 63.

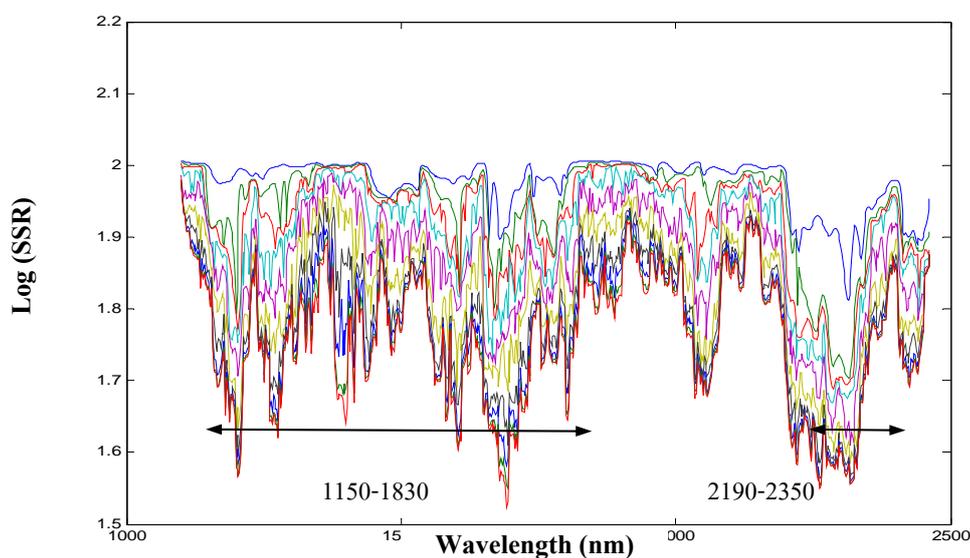


Figure 63 The residue lines obtained by MWPLSR for Pentosan.

The Figure 63 plotted low log (SSR) value in two regions of 1150-1830 nm and 2190-2350 nm. The combination of the absorbance of two regions estimated the pentosan content reported the statistical results in Table 19.

Table 19 The statistical summaries of pentosan calibration model that vary wavelength regions

| No. | Region | Factor | Calibration | | Prediction | | RPD |
|-----|-----------------------|--------|-------------|-------|------------|-------|-------|
| | | | R | SEC* | SEP* | Bias | |
| 1 | 1100-2500 | 3 | 0.782 | 0.820 | 0.795 | 0.262 | 1.308 |
| 2 | 1100-1830 | 1 | 0.388 | 1.213 | 0.923 | 0.212 | 1.127 |
| 3 | 2190-2350 | 1 | 0.602 | 1.052 | 0.772 | 0.224 | 1.347 |
| 4** | 1100-1830 + 2190-2350 | 7 | 0.772 | 0.836 | 0.718 | 0.359 | 1.448 |

* Unit (%W/W), ** Selected Calibration

Both wavelength of absorbance in the two region of 1150-1830 nm and 2190-2350 nm used for making the calibration, the results showed better value than the calibration from the whole wavelength of absorbance. The correlation coefficient of calibration set, SEC, SEP, and RPD were 0.772, 0.836, 0.718, and 1.448 respectively. However, RPD value was low value because standard derivative of prediction set was narrow range. The best calibration was obtained as equation 30.

$$\% \text{ Pentosan} = b_0 + \sum_{i=1100}^{1830} b_i x_i + \sum_{i=2190}^{2350} b_i x_i \quad (30)$$

Where b_0 is y-intercept of regression model, b is regression coefficient at wavelength i , X_i is $d^2 \log(1/R_i)$, i is wavelength

The regression coefficients of equation 30 were several values. Therefore, these values displayed in Figure 64.

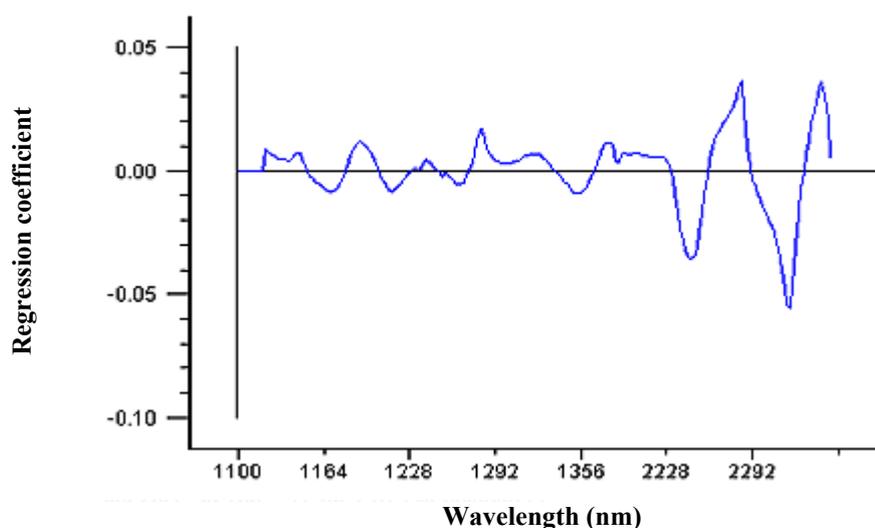


Figure 64 The correlation coefficients of equation 30 with using wavelength in the region of 1100-2500 nm.

The correlation plots between the predicted value and actual value of the calibration set and the validation set illustrated in Figure 65.

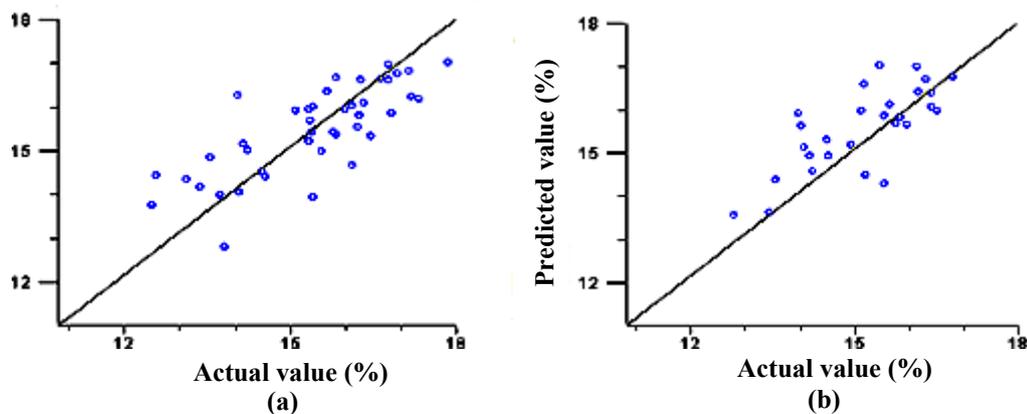


Figure 65 The correlation plots of pentosan between predicted value and the actual value of calibration set (a) and validation set (b).

4.6 Lignin Content

4.6.1 Multiple Linear Regression (MLR) Analysis of Spectral Data and Lignin Content

The statistical summaries were in Appendix Table A7. The correlation of absorbance in the region of 1100-2500 nm and lignin content in each wavelength was pretreated with second derivative to extract important information. The multiple correlations were performed in Figure 66.

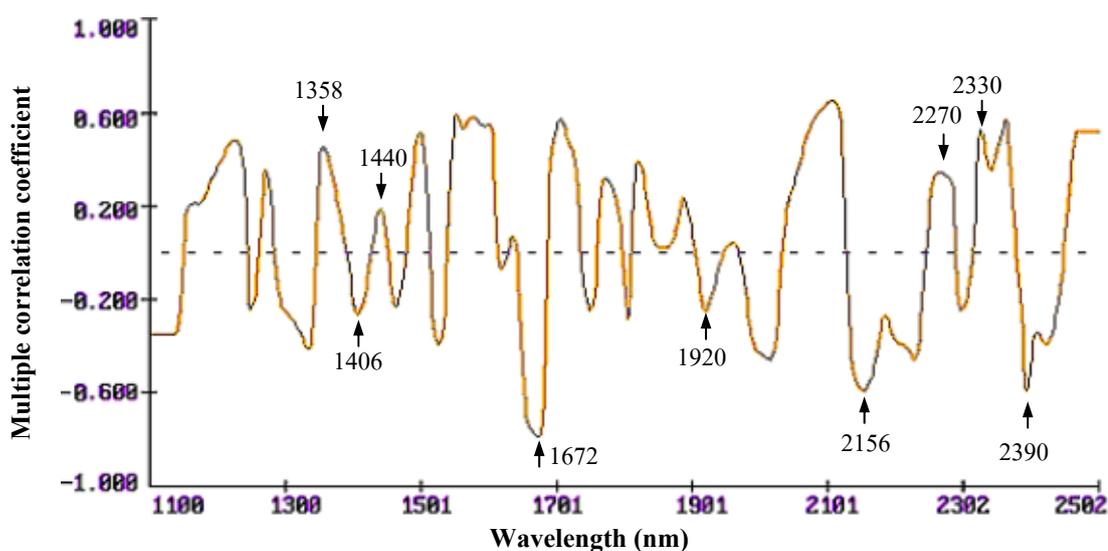


Figure 66 The multiple correlation coefficients of lignin content that pretreated with second derivative.

The spectra were pretreated with second derivative before developed by MLR statistic for lignin prediction. The multiple correlation band that corresponded to the second derivative spectrum of lignin bands were 1358, 1406, 1440, 1672, 1920, 2312, 2156, 2270, 2330, and 2390 nm. The best calibration corresponded absorbance at 1672 nm which related to aromatic peak. The multiple correlation coefficients (R), SEC, SEP, and RPD were 0.833, 0.746, 0.442, and 1.697, respectively. The equation predicted the lignin value was obtained as equation 31.

$$\% \text{ Lignin} = 33.736 - 1556.270 d^2(1/R_{1672}) - 584.836 d^2(1/R_{2430}) \quad (31)$$

The correlations plots between the predicted value and actual value presented in Figure 67.

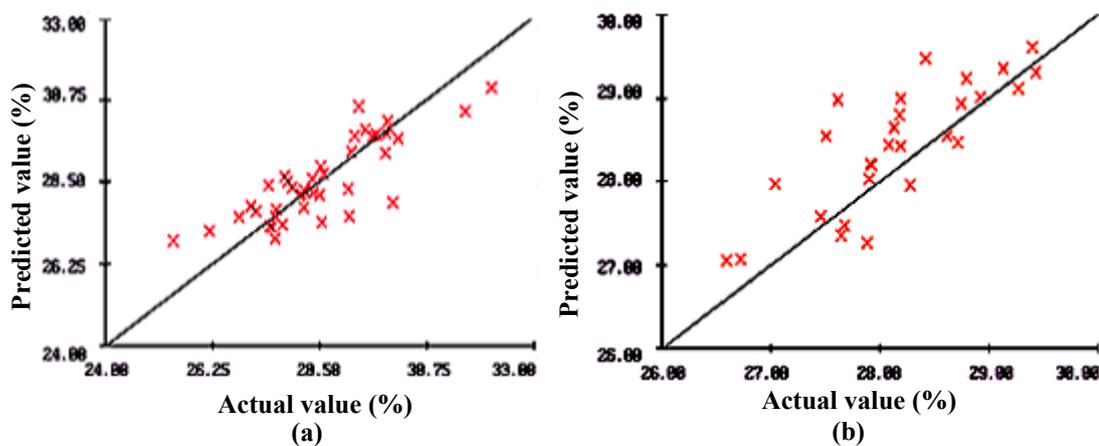


Figure 67 The correlation plot between predicted value and the actual value of calibration set (a) and validation set (b).

4.6.2 Partial Least Square Regression (PLSR) Analysis of Spectral Data and Lignin Content

This analysis compared between different pretreatment before making calibration to estimate the lignin content. The whole wavelength range in the region of 1100-2500 nm was used to calculate and validated as test set method. The statistic summaries reported in Table 20.

Table 20 The statistical summaries of lignin calibration model with different pretreatment methods.

| No. | Pretreatment | Factor | Calibration | | Prediction | | RPD |
|-----|------------------------|--------|-------------|-------|------------|-------|-------|
| | | | R | SEC* | SEP* | Bias | |
| 1 | Average spectra | 9 | 0.851 | 0.695 | 0.593 | 0.325 | 1.265 |
| 2 | MSC | 4 | 0.804 | 0.786 | 0.482 | 0.355 | 1.556 |
| 3 | Second derivatives | 4 | 0.844 | 0.709 | 0.472 | 0.325 | 1.589 |
| 4 | MSC+ second derivative | 4 | 0.851 | 0.694 | 0.472 | 0.292 | 1.589 |

* Unit (%W/W)

From the statistic, the whole wavelength of calibration pretreated by second derivative that provided the similar SEP value as MSC followed by second derivative. Although, the correlation coefficient from previous pretreatment was lower, the SEC and Bias were higher than the another calibration. Therefore, MSC followed by second derivative was selected for pretreatment, then transferred this spectra selected to moving window program. Graph between log (SSR) and wavelengths presented in Figure 68.

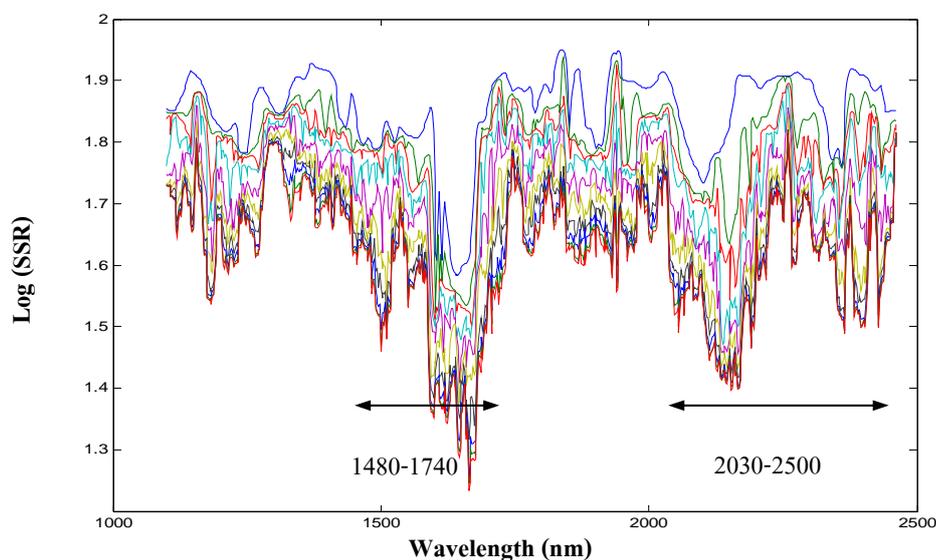


Figure 68 The residue lines obtained by MWPLSR for lignin.

From the Figure 68, absorbance in two regions performed low log (SSR). Absorbances of these two regions try to make calibration. The statistic results were presented in Table 21.

Table 21 The summaries of the statistic in lignin calibration model with vary two wavelength regions

| No. | Region | Factor | Calibration | | Prediction | | RPD |
|-----|-----------------------|--------|-------------|-------|------------|-------|-------|
| | | | R | SEC* | SEP* | Bias | |
| 1 | 1100-2500 | 4 | 0.851 | 0.694 | 0.472 | 0.292 | 1.589 |
| 1 | 1480-1740 | 2 | 0.809 | 0.777 | 0.487 | 0.296 | 1.540 |
| 2 | 2030-2500 | 2 | 0.752 | 0.870 | 0.455 | 0.381 | 1.648 |
| 3** | 1480-1740 + 2030-2500 | 2 | 0.807 | 0.781 | 0.426 | 0.325 | 1.761 |

* Unit (%W/W), ** Selected Calibration

From Table 22, the fit calibration for estimate the lignin contents was used the absorbance in two regions of wavelength. These were the absorbance in the region 1480-1740 nm and 2030-2500 nm. The correlation coefficients (R), SEC, SEP, and RPD were 0.807, 0.781, 0.426, and 1.761, respectively. The best calibration was obtained as equation 32.

$$\% \text{ Lignin} = b_0 + \sum_{i=1480}^{1740} b_i x_i + \sum_{i=2030}^{2500} b_i x_i \quad (32)$$

Where b_0 is y-intercept of regression model, b is regression coefficient at wavelength i , X_i is the absorbance pretreated with MSC and followed by second derivative, i is wavelength. The regression coefficient of this equation showed in Figure 69.

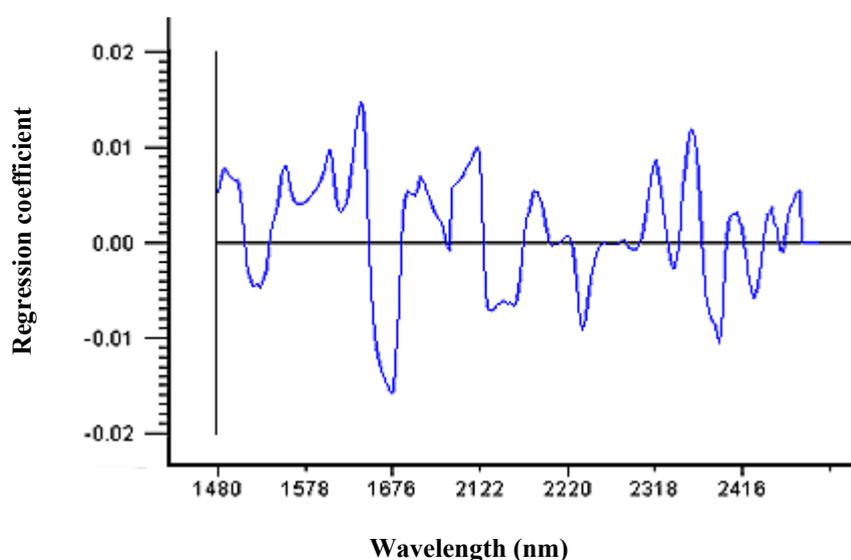


Figure 69 The regression coefficients from equation 32 with using wavelength in the region of 1480-1740 nm and 2030-2500 nm.

The correlation plots between the predicted value and actual value of the calibration set and the validation set were shown as Figure 70.

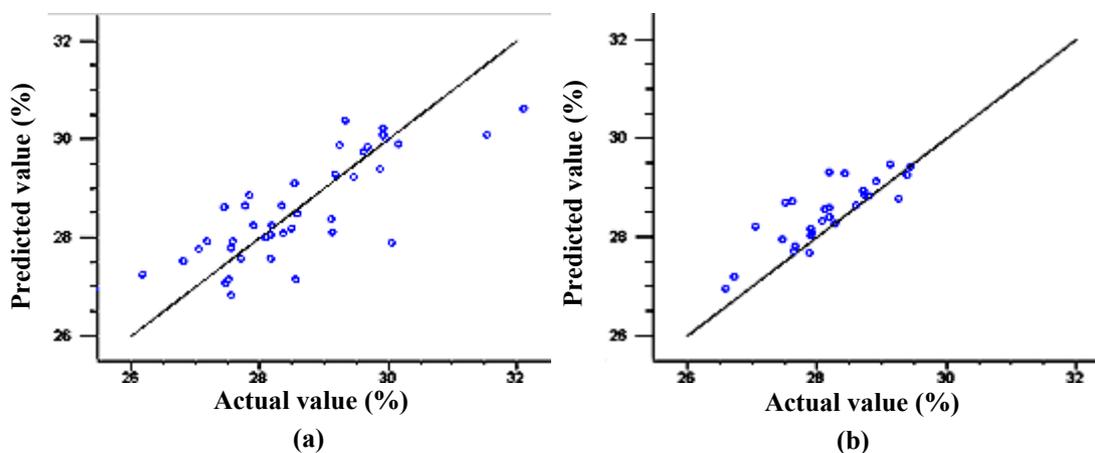


Figure 70 The correlation plots of lignin between predicted value and the measure value of calibration set (a) and validation set (b).

4.7 One percentage of NaOH Solubility

4.7.1 Multiple Linear Regression (MLRR) Analysis of Spectral Data and 1%NaOH Solubility

The second derivative corrections of spectral data were used in order to extract important information from pretreated absorbance data. The correlation between absorbance at in range 1100-2500 nm and 1%NaOH solubility content in each wavelength presented in Figure 71.

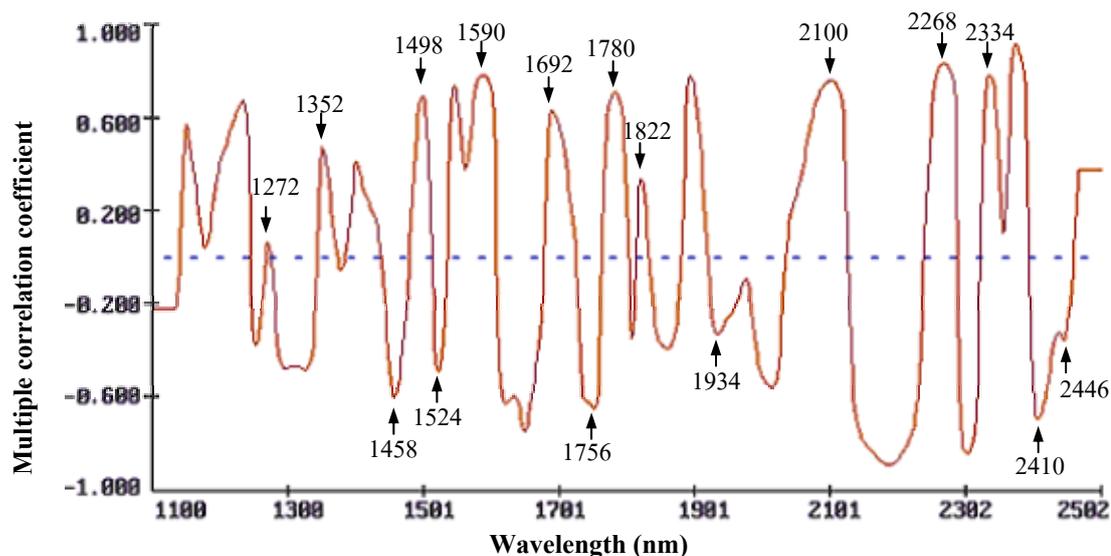


Figure 71 The multiple correlation coefficients of 1% NaOH solubility that pretreated with second derivative.

The selection of wavelength band selected as the nearly second derivative wavelength band of α -cellulose, and xylose standard. All of these were 1272, 1352, 1458, 1498, 1524, 1590, 1692, 1756, 1780, 1822, 1934, 2100, 2268, 2334, 2410, and 2446 nm. The statistic results reported in appendix table A7. The best calibration consist of absorbance at 1934 nm that related to O-H stretching, the other absorbance was 2186 nm, which corresponded to protein band (Burn, 2001). The others band were 1888 and 2316 nm. The multiple correlation coefficient and SEC set were 0.945 and 1.090. The SEP and RPD were 0.044 and 3.704. The calibration equation for 1% NaOH solubility value analysis from NSAS program was obtained as equation 33.

$$1\% \text{NaOH Solubility} = 24.166 + 438.333 d^2 \log(1/R_{1934}) - 2593.623 d^2 \log(1/R_{2186}) + 866.776 d^2 \log(1/R_{1888}) - 895.296 d^2 \log(1/R_{2316}) \quad (33)$$

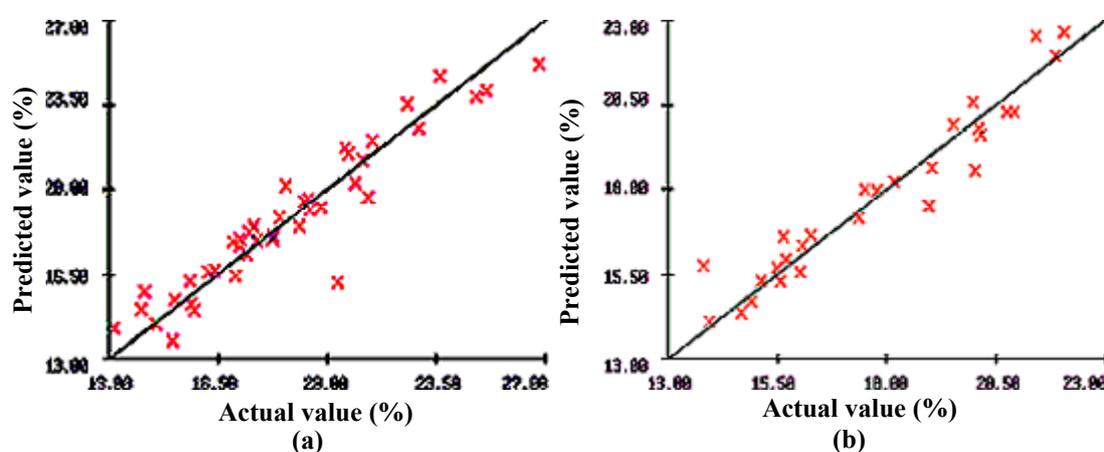


Figure 72 Scatter plot between predicted value and the measure value of calibration set (a) and validation set (b)

4.7.2 Partial Least Square Regression (PLSR) Analysis of Spectral Data and 1%NaOH Solubility

This analysis compared between using the pretreatment data before making calibration for the 1% NaOH solubility value. The whole wavelength in range 1100-2500 nm was calculated and validated as test set method. The statistical results presented in Table 22.

Table 22 The statistical summaries of 1% NaOH solubility value calibration model with pretreatment in vary method

| No. | Pretreatment | Factor | Calibration | | Prediction | | RPD |
|-----|----------------------------|--------|-------------|-------|------------|-------|-------|
| | | | R | SEC* | SEP* | Bias | |
| 1 | Average spectra | 5 | 0.965 | 0.841 | 0.799 | 0.246 | 3.254 |
| 2 | MSC | 7 | 0.976 | 0.696 | 0.597 | 0.075 | 4.355 |
| 3 | Second derivatives | 4 | 0.977 | 0.674 | 0.635 | 0.098 | 4.094 |
| 4 | MSC + second derivative | 3 | 0.975 | 0.705 | 0.596 | 0.046 | 4.362 |

* Unit (%W/W)

The spectra pretreated by MSC and second derivatives performed the best equation. The next step, the pretreated spectra were transferred to moving window program. The result between log (SSR) and wavelength was shown in Figure 73.

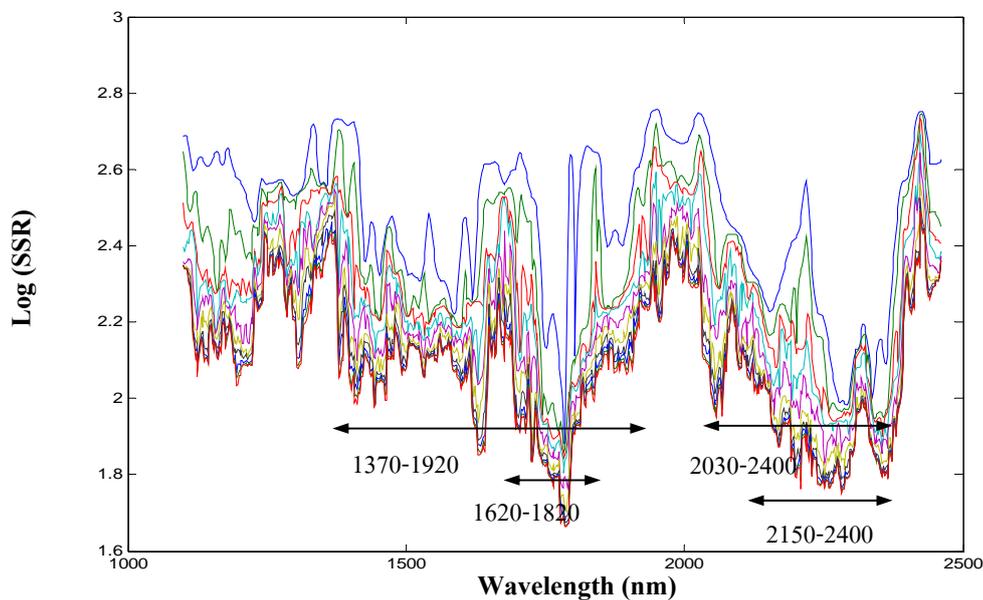


Figure 73 The residue lines obtained by MWPLSR for 1% NaOH solubility.

All four regions of absorbance wavelength; 1370-1920, 1620-1820, 2030-2400, and 2150-2400 nm, were developed the calibration models. The summaries of the statistic were reported in Table 23.

Table 23 Table The statistical summaries of 1% NaOH solubility calibration model with vary wavelength regions.

| No. | Region | Factor | Calibration | | Prediction | | RPD |
|-----|---------------------|--------|-------------|-------|------------|-------|-------|
| | | | R | SEC* | SEP* | Bias | |
| 1 | 1100-2500 | 3 | 0.975 | 0.705 | 0.596 | 0.046 | 4.362 |
| 2 | 1370-1920 | 3 | 0.968 | 0.797 | 0.744 | 0.003 | 3.495 |
| 3 | 2030-2400 | 3 | 0.963 | 0.866 | 0.704 | 0.067 | 3.693 |
| 4 | 1370-1920+2030-2400 | 3 | 0.973 | 0.740 | 0.603 | 0.022 | 4.312 |
| 5 | 1620-1820 | 4 | 0.967 | 0.815 | 0.638 | 0.013 | 4.075 |
| 6 | 2150-2400 | 3 | 0.963 | 0.866 | 0.704 | 0.067 | 3.693 |
| 7** | 1620-1820+2150-2400 | 2 | 0.970 | 0.776 | 0.591 | 0.024 | 4.399 |

* Unit (%W/W), ** Selected Calibration

Both of absorbance wavelengths in the two regions at 1620-1820 nm and 2150-2400 nm used for building the calibration. The statistic result provided better than the whole absorbance wavelength. The correlation coefficients of calibration set, SEC, SEP, and RPD, were 0.970, 0.776, 0.591, and 4.40 respectively. The summaries of equation reported in Table 23, the best calibration was obtained as equation 34.

$$\% \text{ 1\%NaOH solubility} = b_0 + \sum_{i=1620}^{1820} b_i x_i + \sum_{i=2150}^{2400} b_i x_i \quad (34)$$

Where b_0 is y-intercept of regression model, b is regression coefficient at wavelength i , x is $d^2 \log(1/R)$, i is wavelength. There were several values of the regression coefficients for equation 34, which showed in Figure 74.

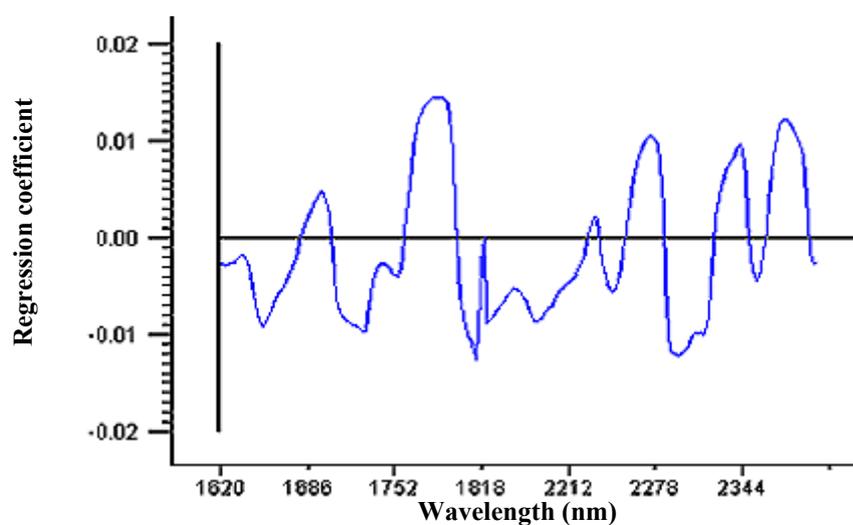


Figure 74 The regression coefficients from equation 34 of 1%NaOH solubility.

The correlation plots between the predicted value of calibration set and prediction set illustrated in Figure 75.

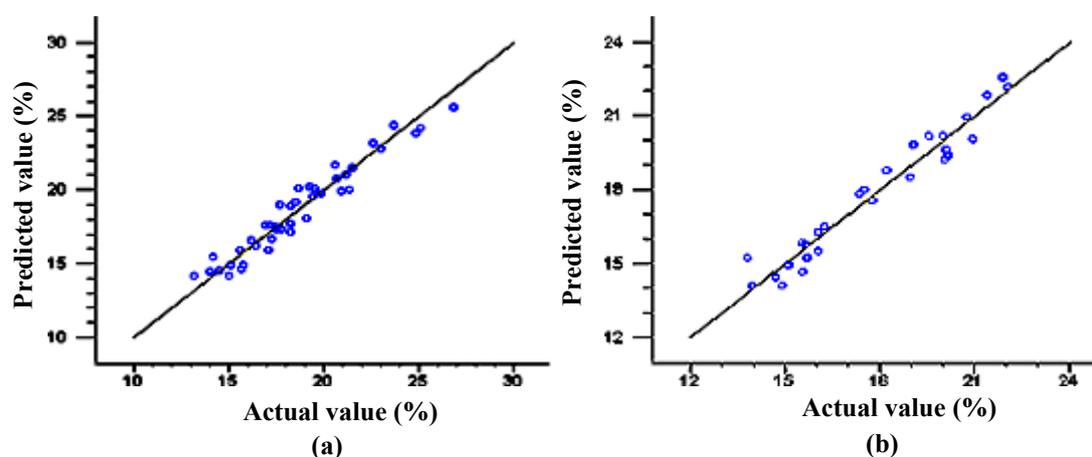


Figure 75 The correlation plots of 1%NaOH solubility between predicted value and actual value of calibration set(a) and validation set(b).

4.8 Ash Content

4.8.1 Multiple Linear Regression (MLR) Analysis of Spectral Data and Ash Content

The correlation between the second derivative absorbance in range 1100-2500 nm and ash in each wavelength presented in Figure 76.

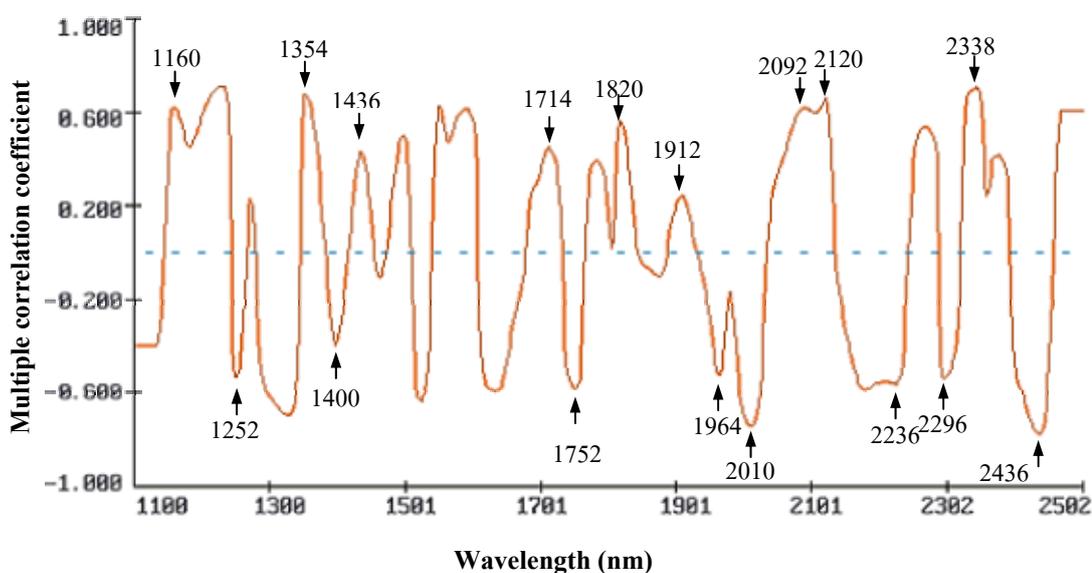


Figure 76 The multiple correlation coefficients of ash content that pretreated with second derivative.

In the theory, ash can not characterize by NIR because ash is inorganic compound. However, the ash was scanned by NIR instrument that presented NIR spectrum as Figure 69. The nearly ash band of multiple linear regression were; 1160, 1252, 1354, 1400, 1436, 1714, 1752, 1820, 1912, 1964, 2010, 2092, 2120, 2236, 2296, 2338, and 2436 nm. All of these bands were selected to calculate in first parameter. The summary of statistic indicated in appendix table A8. The best calibration gave the first absorbance at 1964 nm which was related to O-H stretching and O-H bending of starch and cellulose. The others absorbance were 1296, 1296, 1902 nm. The correlations and SEC were 0.943 and 0.077. The SEP and RPD were 0.063 and 1.905. The calibration for ash value analysis from NSAS program was obtained as equation 35.

$$\% \text{ Ash} = 1.356 - 330.274 d^2 \log(1/R_{1964}) - 394.244 d^2 \log(1/R_{1296}) + 278.757 d^2 \log(1/R_{1974}) - 17.420 d^2 \log(1/R_{1902}) \quad (35)$$

The correlation plots showed between the predicted value and actual value of calibration set and validation set presented in Figure 77.

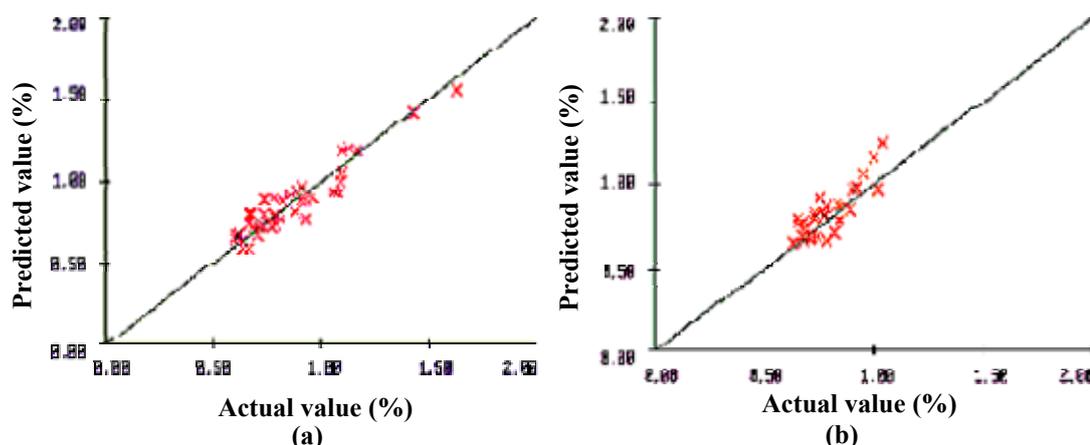


Figure 77 The correlation plots of ash content between predicted value and actual value of calibration set (a) and validation set (b).

4.8.2 Partial Least Square Regression (PLSR) Analysis of Spectral Data and Ash Content

This analysis compared between different pretreatment before making calibration for ash value. The whole wavelength in the region of 1100-2500 nm and validate as test set method. The summary of the statistic result showed in Table 24.

Table 24 The statistical summaries of ash content calibration model with different pretreatment methods.

| No. | Pretreatment | Factor | Calibration | | Prediction | | RPD |
|-----|------------------------|--------|-------------|-------|------------|-------|-------|
| | | | R | SEC* | SEP* | Bias | |
| 1 | Average spectra | 9 | 0.936 | 0.078 | 0.101 | 0.059 | 1.188 |
| 2 | MSC | 5 | 0.864 | 0.111 | 0.101 | 0.071 | 1.188 |
| 3 | Second derivatives | 7 | 0.978 | 0.046 | 0.095 | 0.038 | 1.263 |
| 4 | MSC+ second derivative | 4 | 0.947 | 0.071 | 0.101 | 0.054 | 1.188 |

* Unit (%W/W)

From the result, pretreatment with second derivative gave the best result for making the calibration. The SEC and bias were the lowest value, and then transferred these spectra to moving window program. The result between log (SSR) and wavelength illustrated in Figure 78.

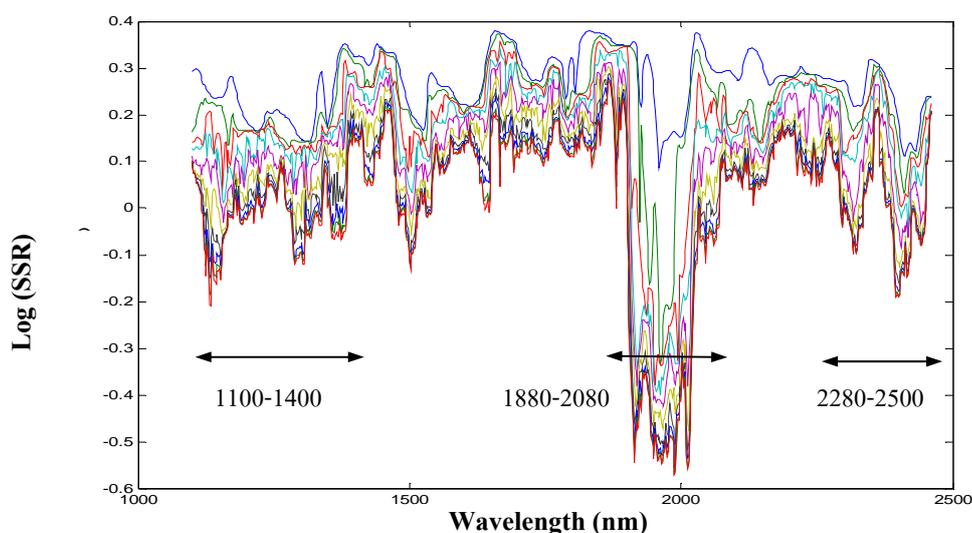


Figure 78 The residue lines obtained by MWPLSR for ash content.

From moving window selection, the three regions in the wavelength range of 1100-1400 nm, 1880-2080 nm were low log (SSR) value. The combination of the absorbance of two and three regions estimated ash content. The summary of statistic results reported in Table 25.

Table 25 The statistical summaries of ash content calibration model that vary wavelength regions

| No. | Region | Factor | Calibration | | Prediction | | RPD |
|-----|--------------------------------------|--------|-------------|-------|------------|-------|-------|
| | | | R | SEC* | SEP* | Bias | |
| 1 | 1100-2500 | 7 | 0.978 | 0.046 | 0.095 | 0.038 | 1.263 |
| 2 | 1880-2080 | 5 | 0.945 | 0.072 | 0.072 | 0.032 | 1.667 |
| 3** | 1100-1400+1880-2080 | 7 | 0.965 | 0.058 | 0.061 | 0.027 | 1.967 |
| 4 | 1100-1400+2280-2500 | 1 | 0.741 | 0.148 | 0.119 | 0.051 | 1.008 |
| 5 | 1880-2080+ 280-2500 | 7 | 0.970 | 0.054 | 0.089 | 0.048 | 1.348 |
| 6 | 1100-1400 + 1880-2080 + 2280-2500 | 8 | 0.982 | 0.042 | 0.097 | 0.027 | 1.237 |

* Unit (%W/W), ** Selected Calibration

The best calibration was used absorbances in range 1100-1400 nm and 1880-2080 nm combination. The correlations and SEC were 0.965 and 0.058. The SEP and RPD are 0.061 and 1.967. The best calibration was obtained as equation 36.

$$\% \text{ Ash} = b_0 + \sum_{i=1100}^{1400} b_i \cdot x_i + \sum_{i=1880}^{2080} b_i \cdot x_i \quad (36)$$

Where b_0 is y-intercept of regression model, b is regression coefficient at wavelength i , X_i is $d^2 \log(1/R)$, i is wavelength. The regression coefficients of this equation presented in Figure 79.

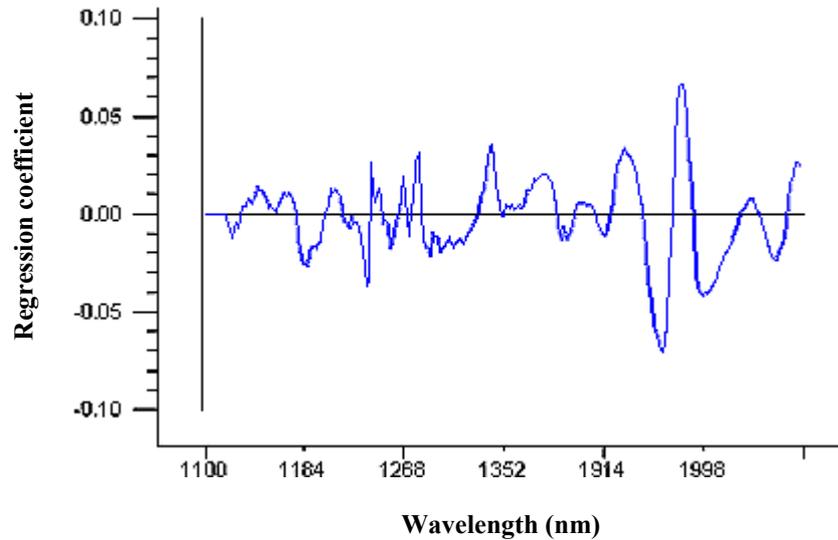


Figure 79 The regression coefficients from equation 36 of ash content with using wavelength 1100-1400 nm and 1880-2080 nm.

The correlation plots between the predicted value and actual value of calibration set and validation set, showed as Figure 80.

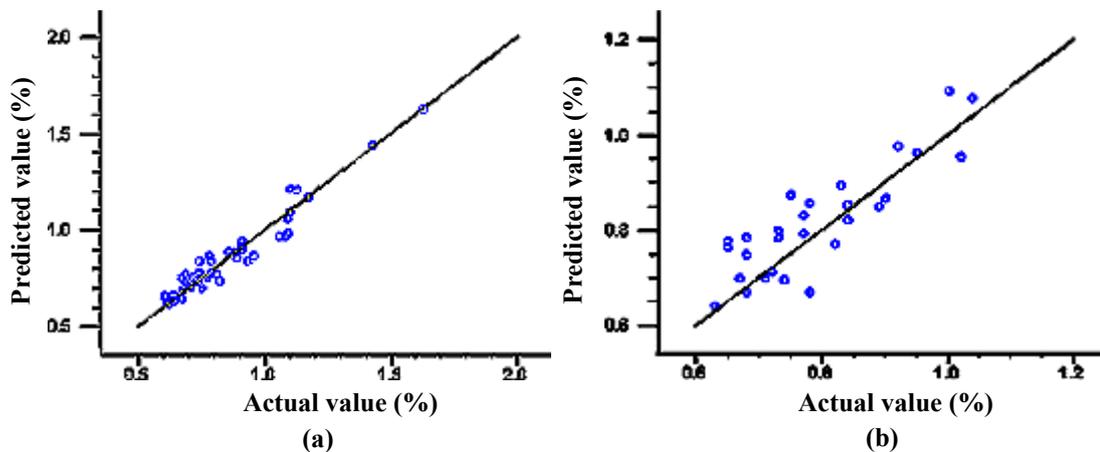


Figure 80 The correlation plots of ash content between predicted value and actual value of calibration set(a) and validation set(b).

4.9 Wood Extractives

4.9.1 Multiple Linear Regression (MLR) Analysis of Spectral Data and Wood Extractives

The correlations of absorbance in the wavelength range 1100-2500 nm and extractive content in each wavelength was pretreated with second derivative to extract important information. The multiple correlations were shown in Figure 81.

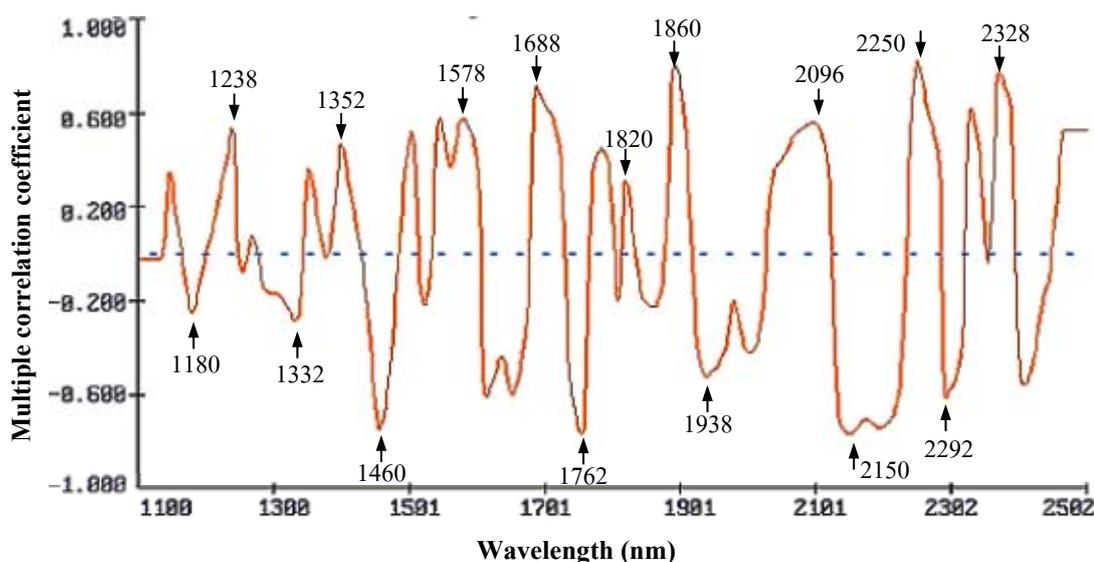


Figure 81 The multiple correlation coefficients of wood extractives that pretreated with second derivative.

The nearly band of multiple linear regression were 1180, 1238, 1332, 1352, 1460, 1580, 1688, 1762, 1820, 1860, 1938, 2096, 2150, 2250, 2292, and 2328 nm. All of these bands were selected to calculate in first parameter. The summary of statistic reported in appendix table A9. The best calibrations gave the first absorbance at 1762 nm that related to 1st overtone of C-H stretching in CH₂, the tentative peak at 2140 nm corresponds C-H stretching combine with C=O stretching or C-H deformation of oil (Burns,2001). The others absorbance were 1708, 1612 nm. The correlations coefficient and standard error of calibration set were 0.925 and 0.384. The standard error and RPD of prediction set were 0.436 and 2.094. The calibration for extractive content analysis from NSAS program was obtained as equation 37.

$$\begin{aligned} \% \text{ Extractive} = & -1.940 - 1830.695 d^2 \log(1/R_{1762}) - 1260.744 d^2 \log(1/R_{2132}) \\ & - 1291.186 d^2 \log(1/R_{1708}) - 1618.670 d^2 \log(1/R_{1612}) \end{aligned} \quad (37)$$

The correlation plots between the predicted value and actual value of calibration set and validation set illustrated in Figure 82.

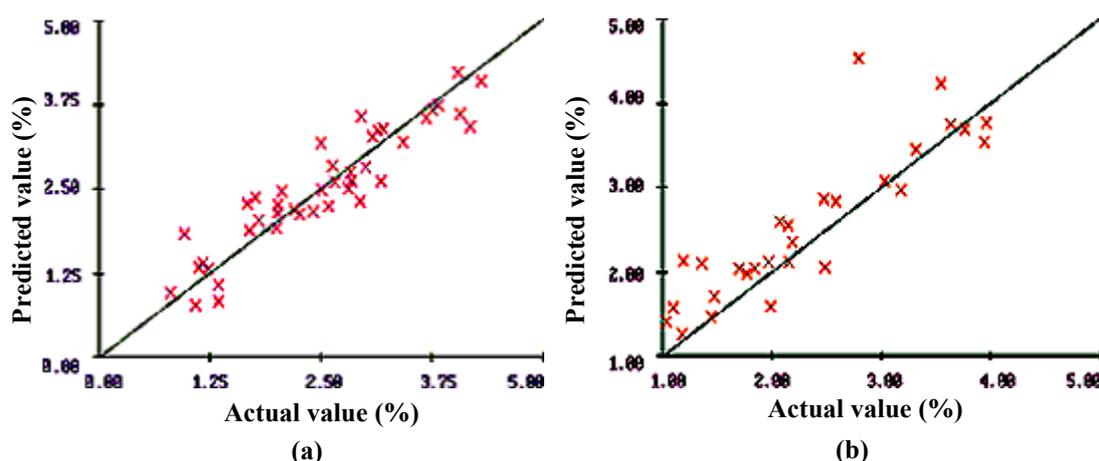


Figure 82 The correlation plots wood extractives between predicted value and actual value of calibration set (a) and validation set (b).

4.9.2 Partial Least Square Regression (PLSR) Analysis of Spectral Data and Wood Extractives

This analysis compared between different pretreatment before making calibration for the extractive content. The whole wavelengths in the range of 1100-2500 nm were calculated and validated as test set method. The statistic summaries indicated in Table 26.

Table 26 The statistical summaries of wood extractives calibration model with different pretreatment methods.

| No. | Pretreatment | Factor | Calibration | | Prediction | | RPD |
|-----|------------------------|--------|-------------|-------|------------|-------|-------|
| | | | R | SEC* | SEP* | Bias | |
| 1 | Average spectra | 2 | 0.957 | 0.281 | 0.522 | 0.233 | 1.749 |
| 2 | MSC | 6 | 0.931 | 0.351 | 0.409 | 0.125 | 2.232 |
| 3 | Second derivatives | 3 | 0.902 | 0.414 | 0.435 | 0.178 | 2.099 |
| 4 | MSC+ second derivative | 8 | 0.971 | 0.229 | 0.399 | 0.144 | 2.288 |

* Unit (%W/W)

From the result, pretreated spectra with MSC and second derivative provided the best result for making the calibration. The SEC and SEP were the lowest value, and then transferred these spectra to moving window program. The result between log(SSR) and wavelength showed as Figure 83.

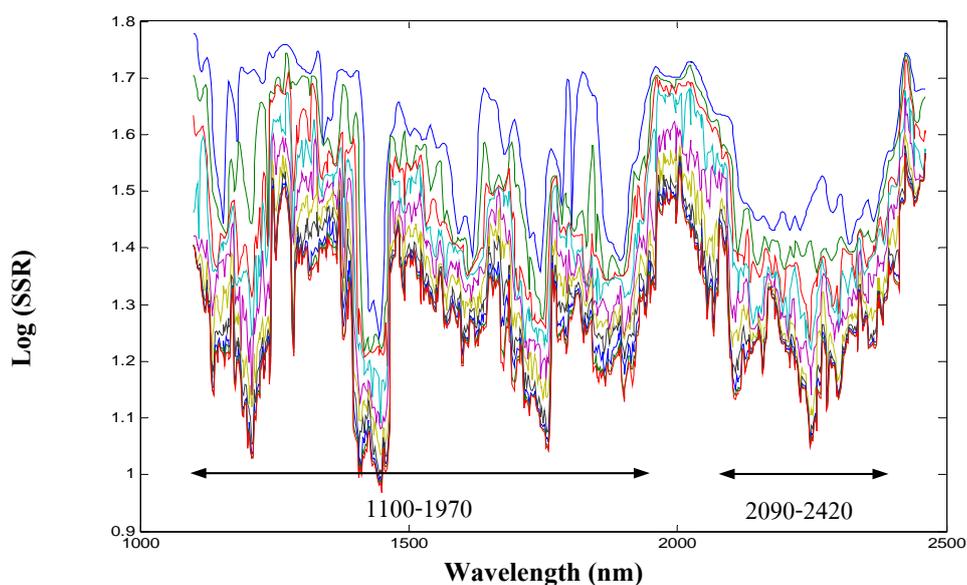


Figure 83 The residue lines obtained by MWPLSR for wood extractives.

From moving window selection, Figure 83, there were two regions in the wavelength of 1100-1970 nm and 2090-2420 nm. The combination of the absorbance of two regions estimated wood extractives. The statistical summaries were reported in Table 27.

Table 27 The statistical summaries in wood extractives calibration model with vary wavelength regions.

| No. | Region | Factor | Calibration | | Prediction | | RPD |
|-----|---------------------|--------|-------------|-------|------------|-------|-------|
| | | | R | SEC* | SEP* | Bias | |
| 1 | 1100-2500 | 8 | 0.971 | 0.229 | 0.399 | 0.144 | 2.288 |
| 2 | 1100-1970 | 7 | 0.963 | 0.258 | 0.381 | 0.166 | 2.396 |
| 3 | 2090-2420 | 7 | 0.963 | 0.258 | 0.504 | 0.191 | 1.812 |
| 4** | 1100-1970+2090-2420 | 7 | 0.967 | 0.246 | 0.378 | 0.175 | 2.415 |

* Unit (%W/W), ** Selected Calibration

The best calibration was used in the wavelength region of 1100-1970 nm combine with 2090-2420 nm. The correlation coefficient of and standard error of calibration set were 0.967 and 0.246. The standard error and RPD of prediction set were 0.378 and 2.415. The best calibration was obtained as equation 38.

$$\% \text{ Extractives} = b_0 + \sum_{i=1100}^{1970} b_i x_i + \sum_{i=2090}^{2420} b_i x_i \quad (38)$$

Where b_0 is y-intercept of regression model, b is regression coefficient at wavelength i , X_i is $d^2 \log(1/R)$, i is wavelength. The regression coefficients of this equation showed in Figure 84.

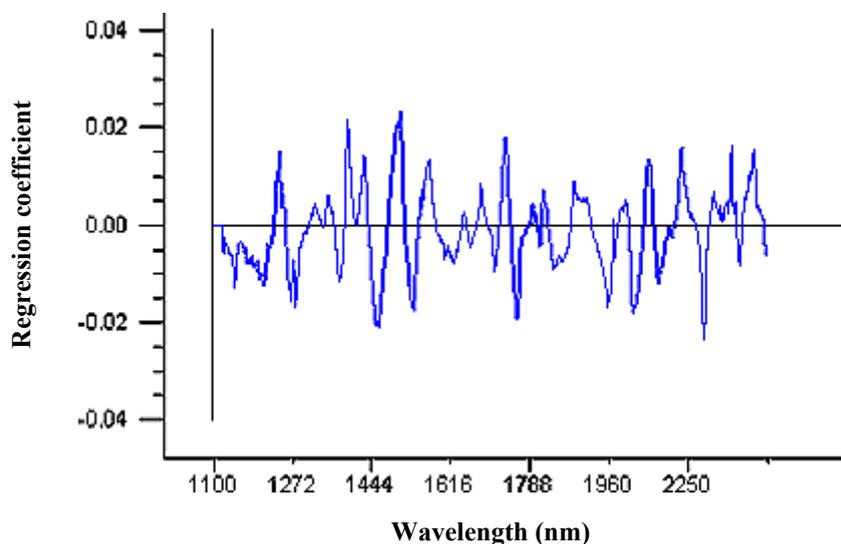


Figure 84 The regression coefficients from equation 38 with using wavelength 1100-1970 nm and 2090-2420 nm.

The correlation plots between the predicted value and actual value of calibration set and validation set presented in Figure 85.

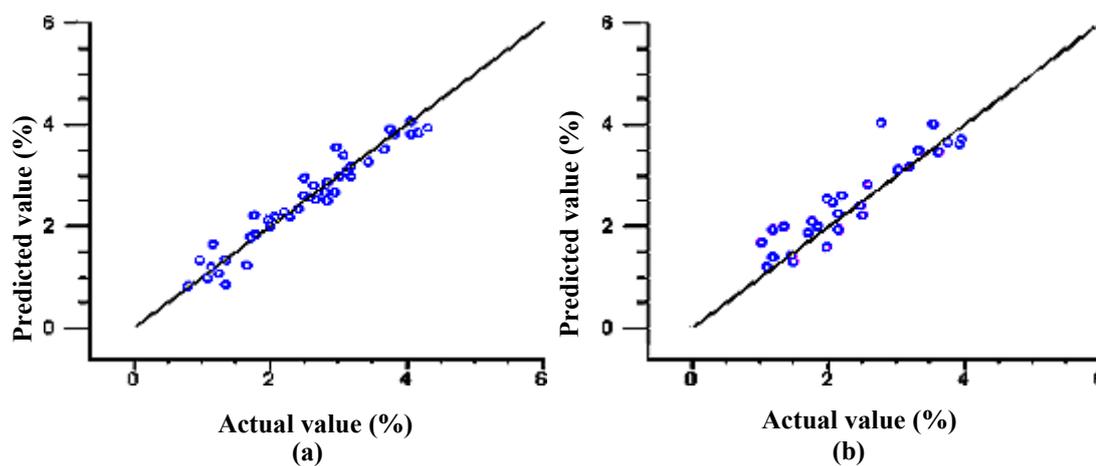


Figure 85 The correlation plots of extractive content between predicted value and actual value of calibration set (a) and validation set (b).

4.10 Glucose Content

4.10.1 Multiple Linear Regression Analysis of Spectral Data and Glucose Content

The correlation of absorbance in range 1100-2500 nm and glucose content in each wavelength was pretreated with second derivative to extract important information. The multiple correlation was shown in Figure 86.

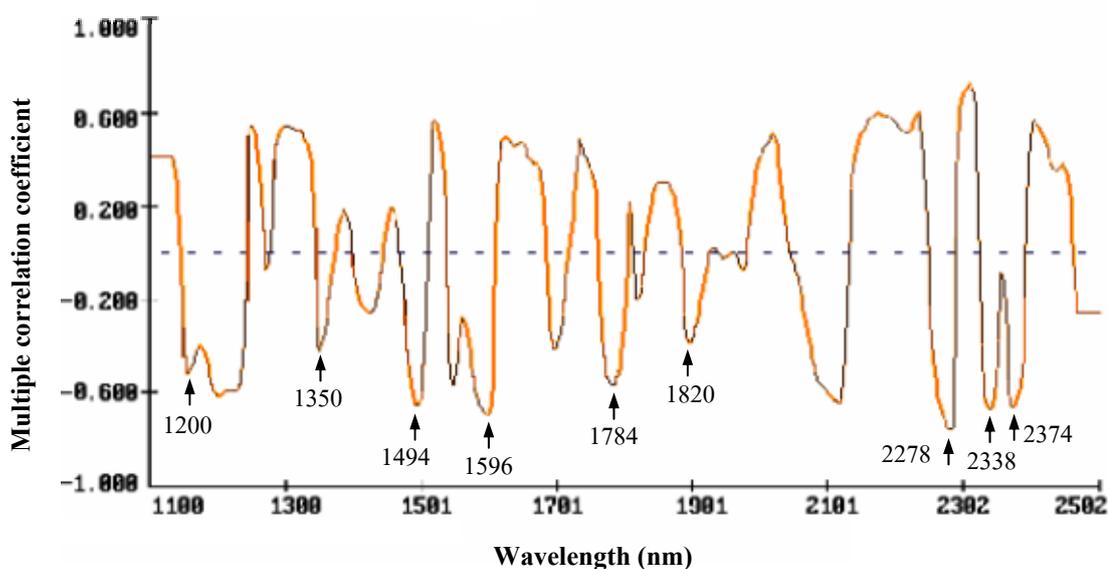


Figure 86 The multiple correlation coefficients of glucose content that pretreated with second derivative.

The automatic selected absorbance at 2278 nm was calculated in first parameter. Moreover, others band of multiple correlation coefficients that related to the second derivative glucose standard spectrum presented in Figure 86, the absorbances at wavelength 1200, 1350, 1494, 1596, 1784, 1820, 2338, 2374 nm were selected to calculate in first parameter. The summary of statistic showed in appendix 10. The best calibrations give the first absorbance at 2278 nm that related to C-H stretching and CH₂ bending combinations or O-H stretching and C-O stretching combination. The others were the wavelengths 1730, 2310, 1814 nm. The correlations coefficient and standard error of calibration set were 0.874 and 1.020. The standard error and RPD of prediction set were 0.827 and 1.753. The Calibration for glucose contents analysis from NSAS program was obtained as equation 39.

$$\% \text{ Glucose} = 35.189 - 1954.682 d^2 \log(1/R_{2278}) + 5273.189 d^2 \log(1/R_{1730}) - 1633.420 d^2 \log(1/R_{2310}) - 4270.591 d^2 \log(1/R_{1814}) \quad (39)$$

The correlation plots between the predicted value and actual value of calibration set and validation set were illustrated in Figure 87.

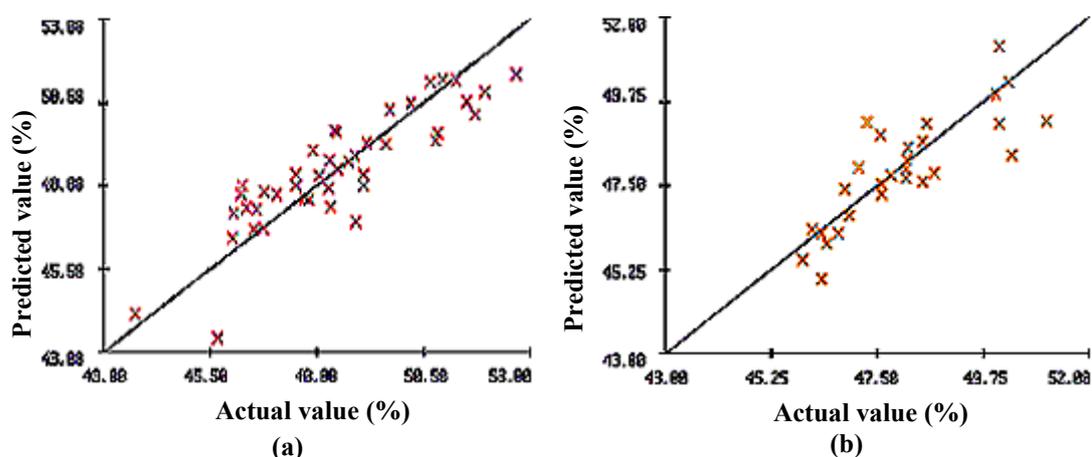


Figure 87 The correlation plots of glucose content between predicted value and actual value of calibration set (a) and validation set (b).

4.10.2 Partial Least Square Regression (PLSR) Analysis of Spectral Data and Glucose Content

This analysis compared between different pretreatment before making calibration for the glucose content. The whole wavelength in the range of 1100-2500 nm was calculated and validated as test set method. The statistic summaries reported in Table 28.

Table 28 The statistical summaries of glucose content calibration model with pretreatment in vary method.

| No. | Pretreatment | Factor | Calibration | | Prediction | | RPD |
|-----|------------------------|--------|-------------|-------|------------|--------|-------|
| | | | R | SEC* | SEP* | Bias | |
| 1 | Average spectra | 8 | 0.874 | 0.967 | 1.010 | -0.650 | 1.436 |
| 2 | MSC | 4 | 0.785 | 1.233 | 0.961 | -0.169 | 1.509 |
| 3 | Second derivatives | 5 | 0.904 | 0.851 | 0.955 | -0.303 | 1.518 |
| 4 | MSC+ second derivative | 6 | 0.921 | 0.680 | 1.123 | 0.266 | 1.291 |

* Unit (%W/W)

From the result, the second derivatives pretreated spectra provided the best statistic result. Although, the comparison of SEP value between the MSC pretreatment and second derivatives were similar, but these second derivative pretreatment gave higher correlation and less SEC value than another pretreatment. The spectra were transferred to moving window program. The result between log (SSR) and wavelength showed in Figure 88.

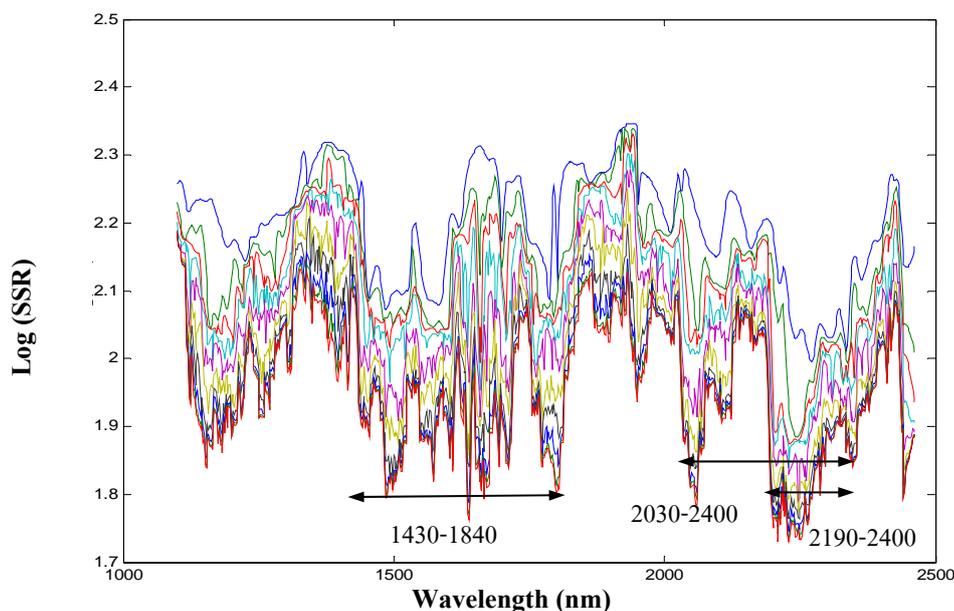


Figure 88 The residue lines obtained by MWPLSR for glucose content.

The three ranges of absorbance used to perform calibrations were 1430-1840, 2030-2400, and 2190-2400 nm. The statistical summaries reported in Table 29.

Table 29 The statistical summaries of glucose content calibration model that vary wavelength regions.

| No. | Region | Factor | Calibration | | Prediction | | RPD |
|-----|---------------------|--------|-------------|-------|------------|--------|-------|
| | | | R | SEC* | SEP* | Bias | |
| 1 | 1430-1840 | 5 | 0.849 | 1.051 | 0.908 | -0.055 | 1.597 |
| 2 | 2190-2400 | 2 | 0.816 | 1.150 | 0.966 | -0.137 | 1.501 |
| 3 | 2030-2400 | 2 | 0.794 | 1.208 | 0.872 | 0.013 | 1.663 |
| 4 | 1430-1840+2190-2400 | 2 | 0.825 | 1.125 | 0.892 | 0.055 | 1.626 |
| 5** | 1430-1840+2030-2400 | 2 | 0.800 | 1.195 | 0.867 | 0.106 | 1.672 |

* Unit (%W/W), ** Selected Calibration

The best calibration was used absorbance in two regions, these were 1430-1840, and 2030-2400 nm for calculation. Although, the correlation coefficient of calibration set were less value, and higher SEC than the once from whole wavelengths, but this calibration give less SEP. The calibrations was

$$\% \text{ Glucose} = b_0 + \sum_{i=1430}^{1840} b_i x_i + \sum_{i=2030}^{2400} b_i x_i \quad (40)$$

Where b_0 is y-intercept of regression model, b is regression coefficient at wavelength i , X_i is $d^2 \log(1/R)$, i is wavelength. The regression coefficients of this equation presented in Figure 89.

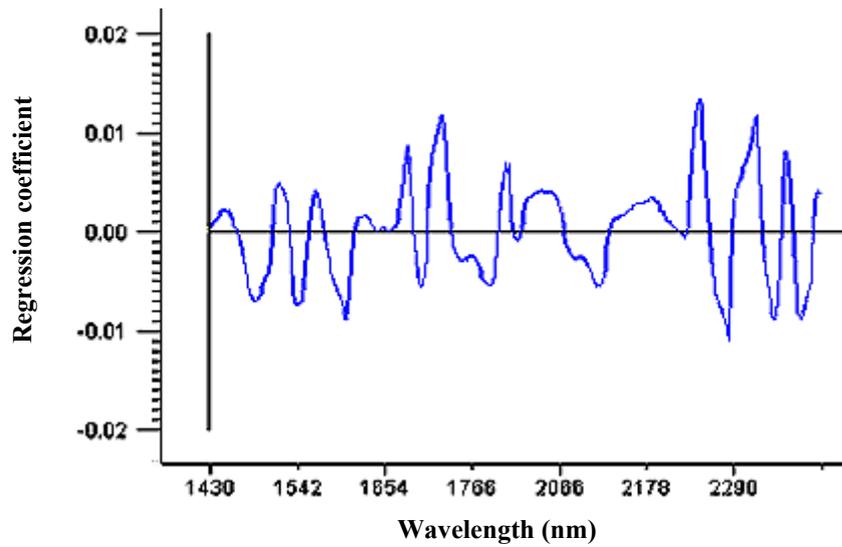


Figure 89 The regression coefficients from equation 40 of glucose with using wavelength 1430-1840 nm.

The correlation plots between the predicted value and actual value of calibration set and validation set were shown as Figure 90.

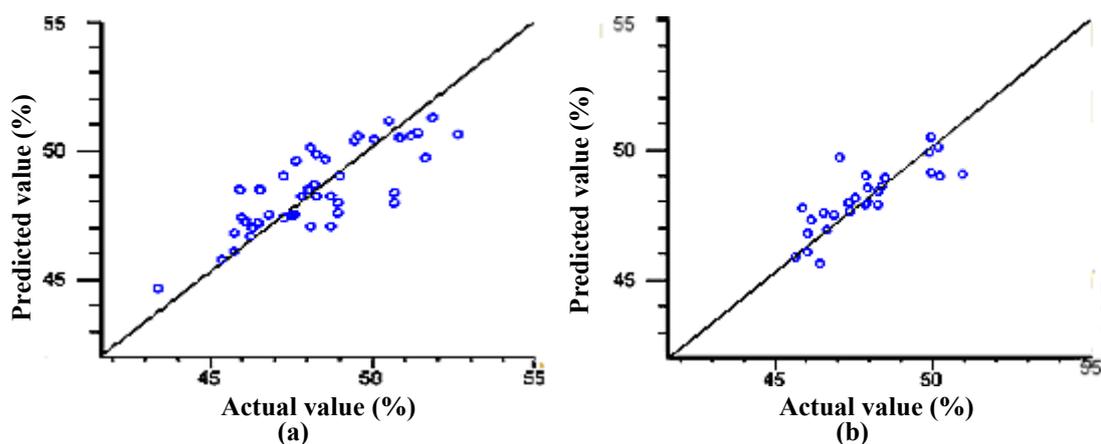


Figure 90 The correlation plots of glucose content between predicted value and actual value of calibration set(a) and validation set(b).

4.11 Xylose Content

4.11.1 Multiple Linear Regression Analysis of Spectral Data and Xylose Content

The correlation of absorbance in the region of 1100-2500 nm and xylose content in each wavelength was pretreated with second derivative to extract important information. The multiple correlations were presented in Figure 91.

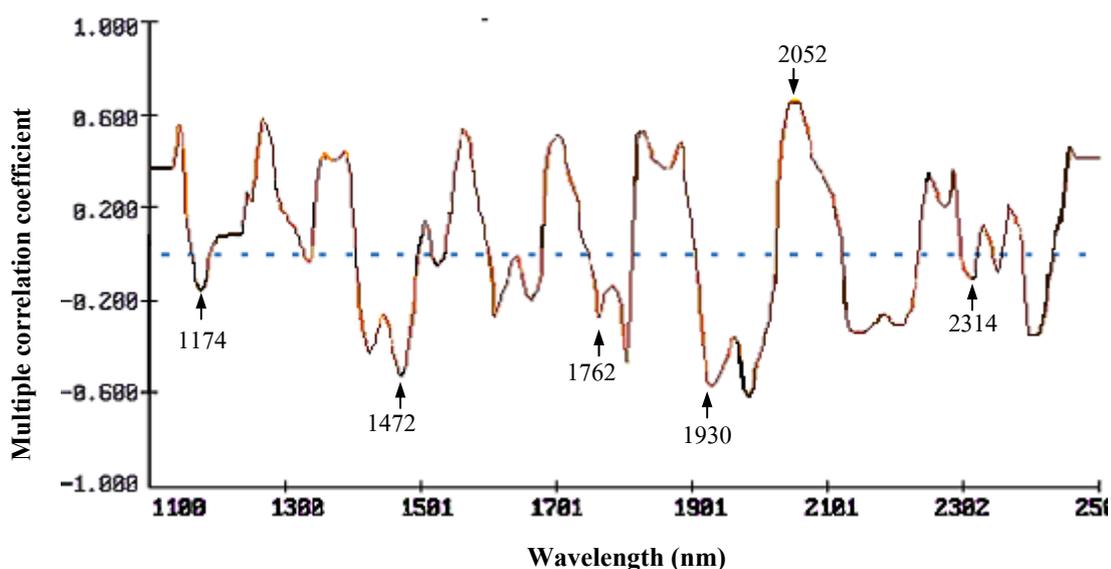


Figure 91 The multiple correlation coefficients of xylose content that pretreated with second derivative.

The first absorbance by automatic selection was the absorbance at 2052 nm. The others selected at absorbance were 1174, 1472, 1762, 1930, 2314 nm. The statistic result indicated in appendix table A11. The calibration reported lowest SEP that consists of absorbance 1762 nm for the first parameter. The second and third absorbance from automatic selection were 2048, 1144 nm. Burn (2001) reported the tentative band at 1765 nm that relate to C-H stretch 1st overtone of CH₂. The absorbance at 2084 related to xylan band (Shimleck, 1996), and 1144 relate to C-H second overtone. The correlation coefficient and standard error of calibration set were 0.736 and 1.660. The standard error and RPD of prediction set were 1.210 and 1.802. The calibration for xylose content evaluation from NSAS program was obtained as equation 41.

$$\begin{aligned} \% \text{ Xylose} = & -3.391 + 5051.535 d^2 \log(1/R_{1762}) + 1310.942 d^2 \log(1/R_{2084}) \\ & - 2445.550 d^2 \log(1/R_{1144}) \end{aligned} \quad (41)$$

The correlation plots between the predicted value and actual value of calibration set and validation set presented in Figure 92.

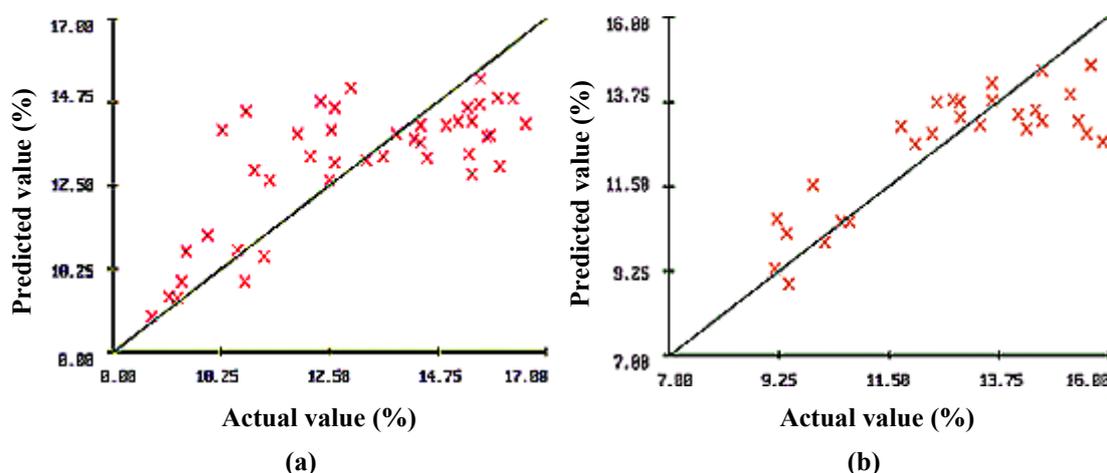


Figure 92 The correlation plots of xylose content between predicted value and actual value of calibration set (a) and validation set (b).

4.11.2 Partial Least Square Regression (PLSR) Analysis of Spectral Data and Xylose Content

This analysis compared between different pretreatment before making calibration to estimate xylose content. The whole wavelength in the region of 1100-2500 nm was used to calibrated and validated as test set method. The statistical summaries were reported in Table 30.

Table 30 The statistical summaries of xylose content calibration model with different pretreatment methods.

| No. | Pretreatment | Factor | Calibration | | Prediction | | RPD |
|-----|-------------------------|--------|-------------|-------|------------|-------|-------|
| | | | R | SEC* | SEP* | Bias | |
| 1 | Average spectra | 3 | 0.658 | 1.788 | 1.386 | 0.135 | 1.573 |
| 2 | MSC | 2 | 0.654 | 1.796 | 1.372 | 0.170 | 1.589 |
| 3 | Second derivatives | 2 | 0.674 | 1.753 | 1.362 | 1.362 | 1.601 |
| 4 | MSC + second derivative | 3 | 0.777 | 1.494 | 1.534 | 0.079 | 1.421 |

* Unit (%W/W)

The second derivative spectra reported the lowest SEP, and then transferred these spectra to moving window program. The result between log (SSR) and wavelength showed in Figure 93.

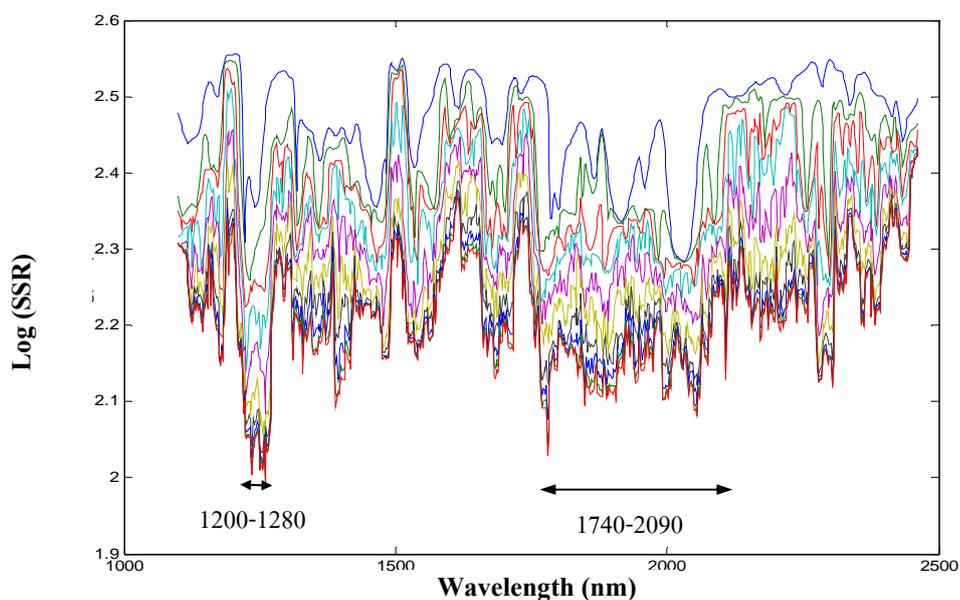


Figure 93 The residue lines obtained by MWPLSR for xylose content.

The residual line character was difficult for selection. However, two absorbance regions in range 1200-1280 nm, and 1740-2090 nm were selected to establish calibration. The statistical summaries were written in Table 31.

Table 31 The statistical summaries of xylose content calibration model with vary wavelength regions.

| No. | Region | Factor | Calibration | | Prediction | | RPD |
|-----|-----------------------|--------|-------------|-------|------------|-------|-------|
| | | | R | SEC* | SEP* | Bias | |
| 1 | 1100-2500 | 2 | 0.674 | 1.753 | 1.362 | 1.362 | 1.601 |
| 2 | 1200-1280 | 2 | 0.627 | 1.850 | 1.407 | 0.225 | 1.549 |
| 3 | 1740-2090 | 2 | 0.732 | 1.616 | 1.363 | 0.128 | 1.599 |
| 4** | 1200-1280 + 1740-2090 | 2 | 0.720 | 1.648 | 1.280 | 0.205 | 1.703 |

* Unit (%W/W), ** Selected Calibration

The absorbance spectra at 1200-1820 nm and 1740-2090 nm were combined and they gave better result than the whole regions for calibrations. The correlation coefficients of calibration set, SEP, and RPD, were 0.720, 1.280, and 1.703 respectively. The best calibration was obtained as equation 42.

$$\% \text{ Xylose} = b_0 + \sum_{i=1200}^{1280} b_i x_i + \sum_{i=1740}^{2090} b_i x_i \quad (42)$$

Where b_0 is y-intercept of regression model, b is regression coefficient at wavelength i , X_i is the absorbance that was pretreated with MSC and second derivative, i is wavelength. The regression coefficients of this equation were illustrated in Figure 94.

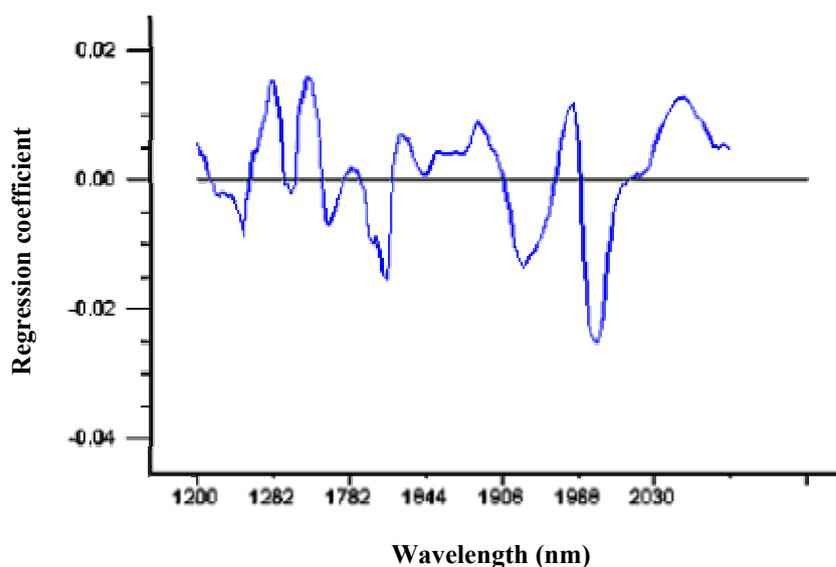


Figure 94 The regression coefficients of xylose content from equation 42 with using wavelength in the region 1200-1280 nm and 1740-2090 nm.

The correlation plots between the predicted value and actual value of calibration set and validation set were presented in Figure 95.

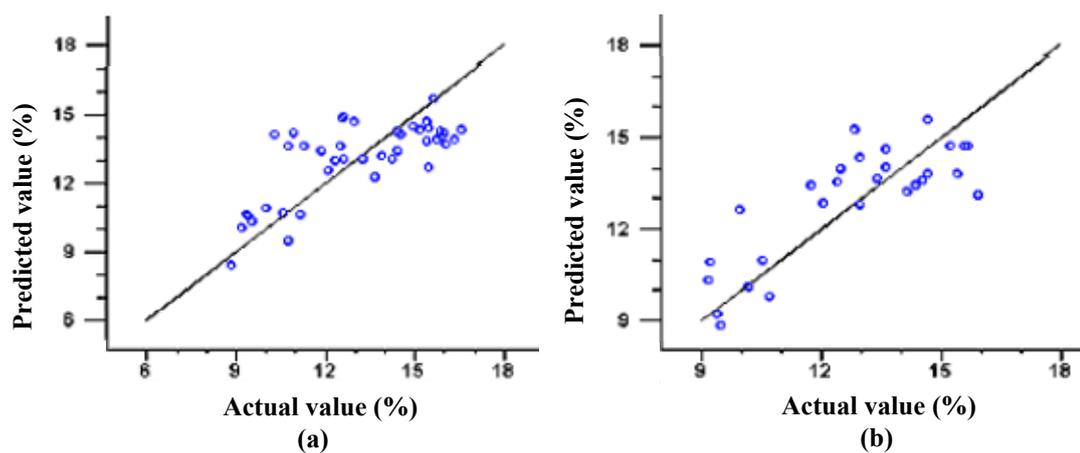


Figure 95 The correlation plots of xylose content between predicted value and actual value of calibration set (a) and validation set (b).

4.12 Arabinose Content

Arabinose was a little component. Mostly arabinose was stayed in arabinogalactan and arabinan form that was the ones of components in hemicellulose (Alen, 2000).

4.12.1 Multiple Linear Regression (MLR) Analysis of Spectral Data and Arabinose Content

The correlation of absorbance in the wavelength range at 1100-2500 nm and arabinose content in each wavelength was pretreated with second derivative to extract important information. The multiple correlations showed in Figure 96.

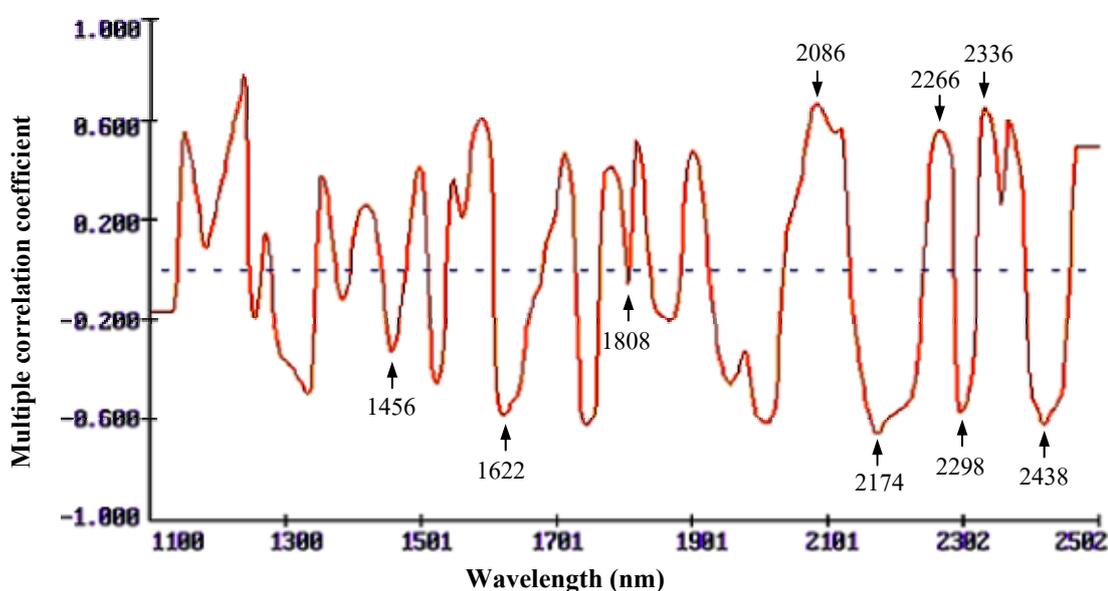


Figure 96 The multiple correlation coefficients of arabinose content that pretreated with second derivative.

The first absorbance selection was selected from the correlations correspond near the second derivative band of arabinose standard. The selected first absorbances were 1456, 1622, 1808, 2086, 2174, 2266, 2298, 2336, and 2438 nm. Moreover the automatic selection was 1238 nm, the statistic result were reported in appendix table A12. The calibration performed the lowest value which consisted of absorbance at 2266 nm for the first parameter. The second and third absorbance were 1240, 1152 nm which was from automatic selection. Burn (2001) reported the tentative band at 2270 nm that relate to O-H stretching combine C-O stretching. The band at 1240 nm related to 1st of O-H stretching and the C-O stretching combination. The tentative band at 1150 nm related to 2nd of C-H stretching (Weyer, 2002). The correlation coefficient and standard error of calibration set were 0.843 and 0.095. The standard error and RPD of prediction set were 0.086 and 1.628. The calibration for arabinose content analysis from NSAS program was obtained as equation 43.

$$\begin{aligned} \% \text{Arabinose} = & -0.212 + 9.626 d^2 \log(1/R_{2266}) + 1226.015 d^2 \log(1/R_{1240}) \\ & + 115.150 d^2 \log(1/R_{1152}) \end{aligned} \quad (43)$$

The correlation plots between the predicted value and actual value of calibration set and validation set were presented in Figure 97.

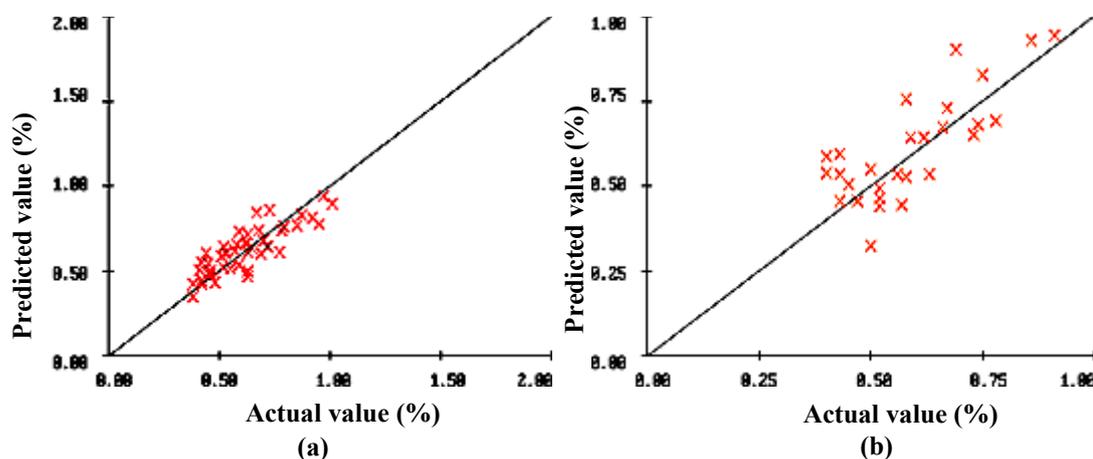


Figure 97 The correlation plots of arabinose content between predicted value and actual value of calibration set (a) and validation set (b).

4.12.2 Partial Least Square Regression (PLSR) Analysis of Spectral Data and Arabinose Content

This analysis compared between the different pretreatment before making calibration was for arabinose content. The whole wavelength in the range of 1100-2500 nm was calculated and validated as test set method. The statistical summaries reported in Table 32.

Table 32 The statistical summaries of arabinose content calibration model with different pretreatment methods.

| No. | Pretreatment | Factor | Calibration | | Prediction | | RPD |
|-----|------------------------|--------|-------------|-------|------------|-------|-------|
| | | | R | SEC* | SEP* | Bias | |
| 1 | Average spectra | 3 | 0.663 | 0.127 | 0.099 | 0.021 | 1.414 |
| 2 | MSC | 1 | 0.605 | 0.135 | 0.092 | 0.022 | 1.522 |
| 3 | Second derivatives | 2 | 0.750 | 0.112 | 0.104 | 0.014 | 1.346 |
| 4 | MSC+ second derivative | 2 | 0.783 | 0.106 | 0.092 | 0.012 | 1.522 |

* Unit (%W/W)

Although, the calibration from average spectra and MSC and second derivative pretreatment spectra gave equal SEP, but the SEC value from MSC and second derivative pretreatment spectra was lower value. The correlation coefficient was higher value. Therefore, this spectral pretreatment was selected for made the calibration. The spectra were transferred to moving window program. The result between log (SSR) and wavelength was illustrated in Figure 98.

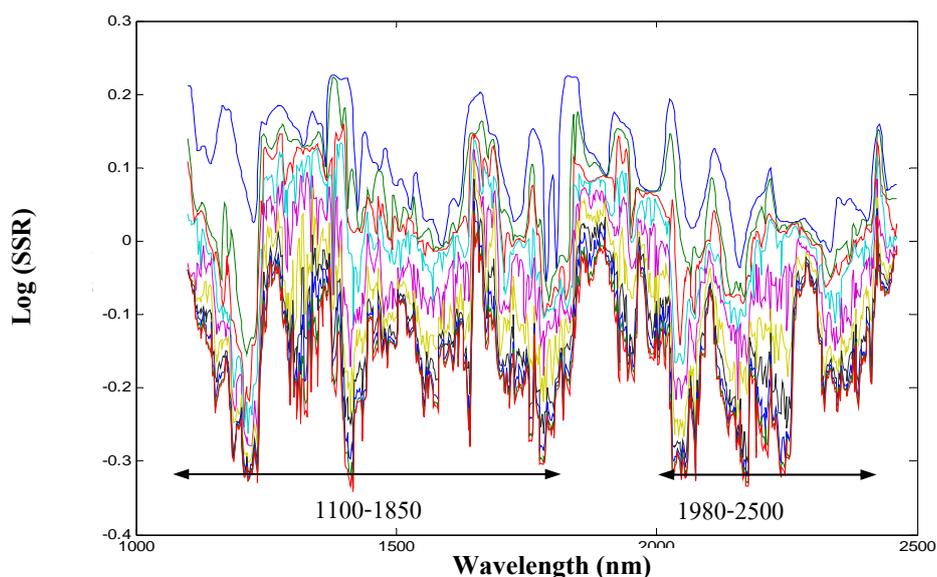


Figure 98 The residue lines obtained by MWPLSR for arabinose content.

From Figure 98, it showed the similar log (SSR) value. However, there was a narrow range give higher log (SSR) than the other region. The residual lines plotted in two regions that showed log (SSR) value. The region in the range of 1100-1850 nm and 1980-2500 nm shows the lowest log (SSR) value. The summaries of statistic result showed in Table 33.

Table 33 The statistical summaries of arabinose content calibration model that vary wavelength regions.

| No. | Region | Factor | Calibration | | Prediction | | RPD |
|-----|-----------------------|--------|-------------|-------|------------|-------|-------|
| | | | R | SEC* | SEP* | Bias | |
| 1 | 1100-2500 | 2 | 0.783 | 0.106 | 0.092 | 0.012 | 1.522 |
| 2 | 1100-1850 | 2 | 0.765 | 0.109 | 0.093 | 0.015 | 1.505 |
| 3 | 1980-2500 | 2 | 0.739 | 0.114 | 0.093 | 0.007 | 1.505 |
| 4 | 1100-1850 + 1980-2500 | 2 | 0.781 | 0.106 | 0.095 | 0.012 | 1.474 |

* Unit (%W/W)

From Table 32, the calibration was used the absorbance in the whole wavelength absorbance gave the highest the correlation coefficient for compared the different regions. The correlation coefficient and standard error of calibration set were 0.783 and 0.106, the standard error of prediction and RPD were 0.092, and 1.522, respectively. The best calibration was obtained as equation 44.

$$\% \text{ Arabinose} = b_0 + \sum_{i=1100}^{2500} b_i x_i \quad (44)$$

Where b_0 is y-intercept of regression model, b is regression coefficient at wavelength i , X_i is the absorbance that was pretreated with MSC and second derivative, i is wavelength. The regression coefficient of this equation was illustrated in Figure 99.

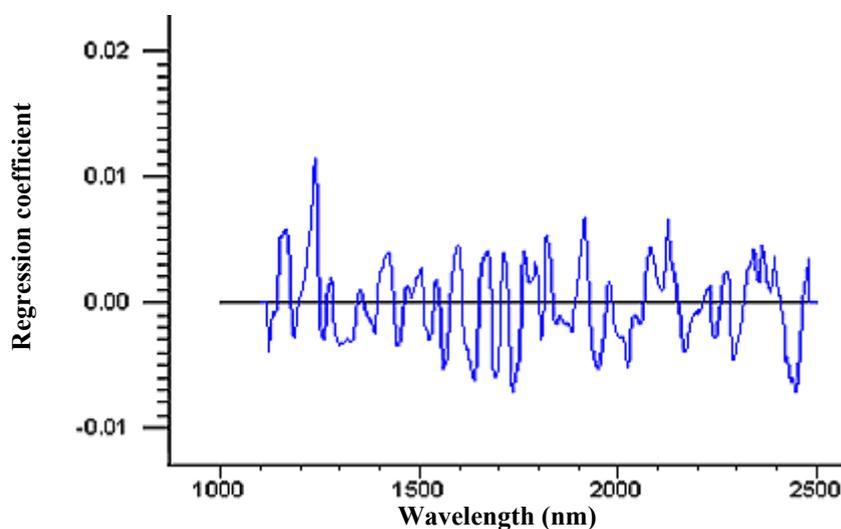


Figure 99 The regression coefficients of arabinose from equation 44 with using wavelength in the region of 1100-2500 nm.

The correlation plots between the predicted value and actual value of calibration set and validation set were shown in Figure 100.

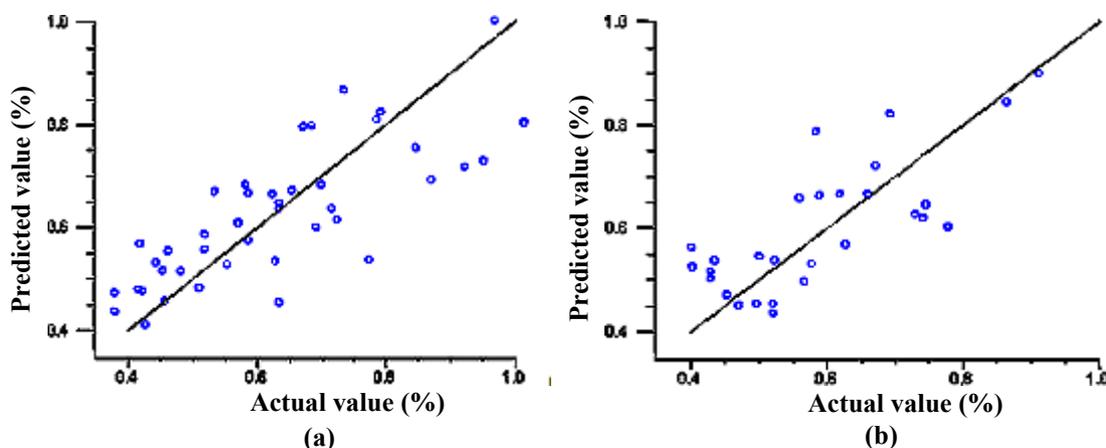


Figure 100 The correlation plots of arabinose content between predicted value and actual value of calibration set (a) and validation set (b).

4.13 Galactose Content

Galactose was consisted in arabinogalactan and galactans form that was one of the components in hemicellulose (Alen, 2000).

4.13.1 Multiple Linear Regression (MLR) Analysis of Spectral Data and Galactose Content

Each absorbance spectra were pretreated second derivative to extract important information before calculation. Figure 101 showed the multiple correlation coefficients that derived from the analysis between the pretreated spectra galactose content.

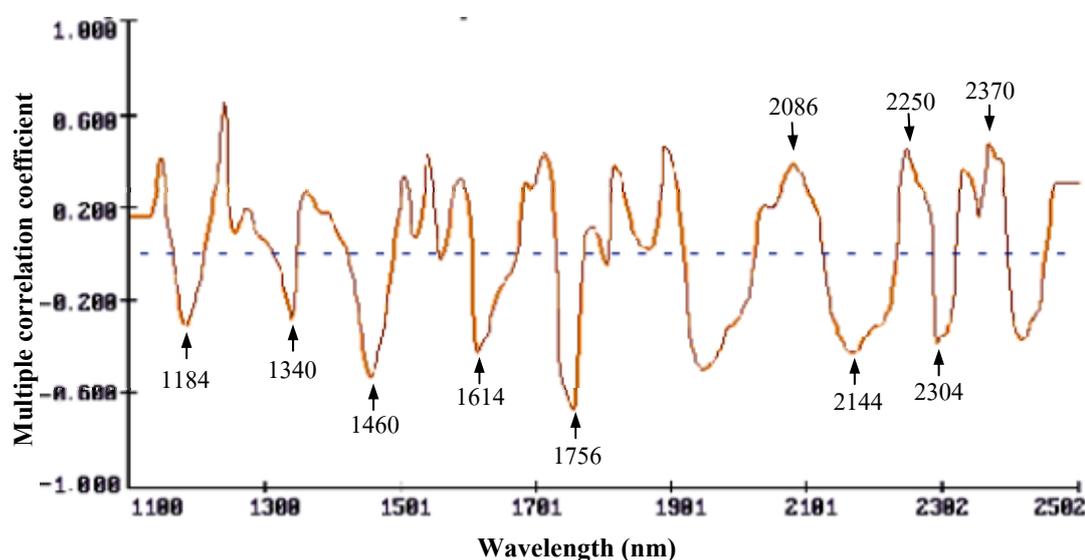


Figure 101 The multiple correlation coefficients of galactose content that pretreated with second derivative.

The first absorbance selection was selected from the correlations correspond near the second derivative band of galactose standard. The automatic selection of first absorbance was 1756 nm and the manual selection absorbance were 1184, 1340, 1460, 1614, 1756, 2086, 2144, 2250, 2304, and 2370 nm. The calibration gave the lowest value consist of absorbance at 2144 nm for the first parameter. The second and third absorbance were 1754, 2200 nm which was from automatic selection. Burn(2001) reported the tentative band at 2170 nm that relate to C-H stretching combine C=O stretching. The band at 1756 nm has not been report, however Weyer (2002) reported in which the 1st overtone of C-H stretching occurred between bands at 1700-1800 nm. The band at 2200 nm correspond to C-H stretching and C=O stretching combination. The correlation coefficient and standard error of calibration set were 0.793 and 0.334. The standard error of prediction and RPD are 0.206 and 1.942. The calibration for galactose content estimation from NSAS program was obtained as equation 45.

$$\begin{aligned} \% \text{ Galactose} = & -0.044 - 347.665 d^2 \log(1/R_{2144}) - 2063.784 d^2 \log(1/R_{1754}) \\ & + 359.644 d^2 \log(1/R_{2200}) \end{aligned} \quad (45)$$

The correlation plots between the predicted value and actual value of calibration set and validation set was presented in Figure 102.

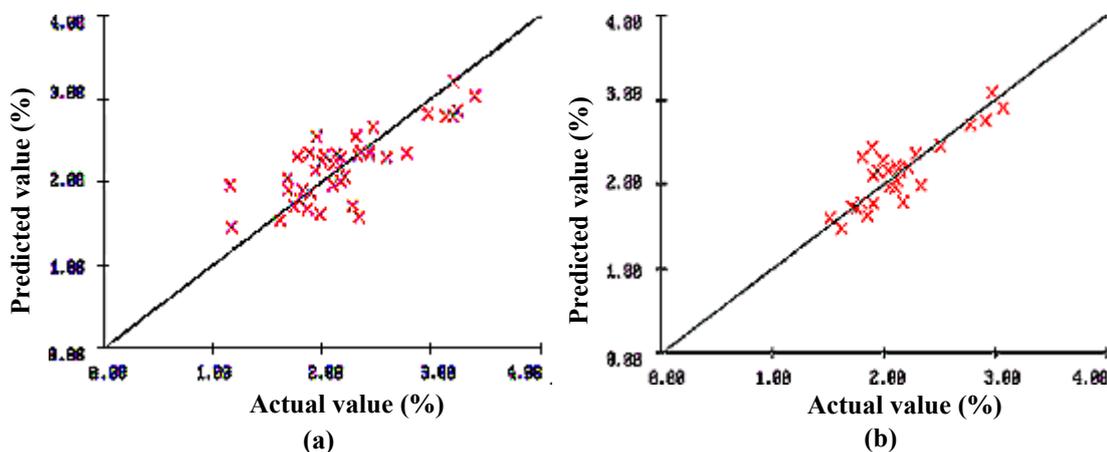


Figure 102 The correlation plots of galactose content between predicted value and actual value of calibration set (a) and validation set (b).

4.13.2 Partial Least Square Regression (PLSR) Analysis of Spectral Data and Galactose Content

This analysis compared between different pretreatment before making calibration to estimate the galactose content. The whole wavelength in the region of 1100-2500 nm was used to calculate and validate as test set method. The statistical summaries of the result showed in Table 34.

Table 34 The statistical summaries of galactose content calibration model with different pretreatment methods.

| No. | Pretreatment | Factor | Calibration | | Prediction | | RPD |
|-----|---------------------------|--------|-------------|-------|------------|--------|-------|
| | | | R | SEC* | SEP* | Bias | |
| 1 | Average spectra | 6 | 0.718 | 0.395 | 0.265 | 0.008 | 1.509 |
| 2 | MSC | 5 | 0.749 | 0.348 | 0.247 | -0.011 | 1.619 |
| 3 | Second derivatives | 3 | 0.756 | 0.345 | 0.206 | 0.038 | 1.942 |
| 4 | MSC+ second derivative | 1 | 0.809 | 0.310 | 0.201 | 0.016 | 1.990 |

* Unit (%W/W)

The fit calibration was performed the pretreated with MSC and second derivative spectra that gave the lowest SEC value and SEP value, but the highest correlation value. The spectra were transferred to moving window program to selected wavelength that related galactose value. The result between log (SSR) and wavelength showed in Figure 103.

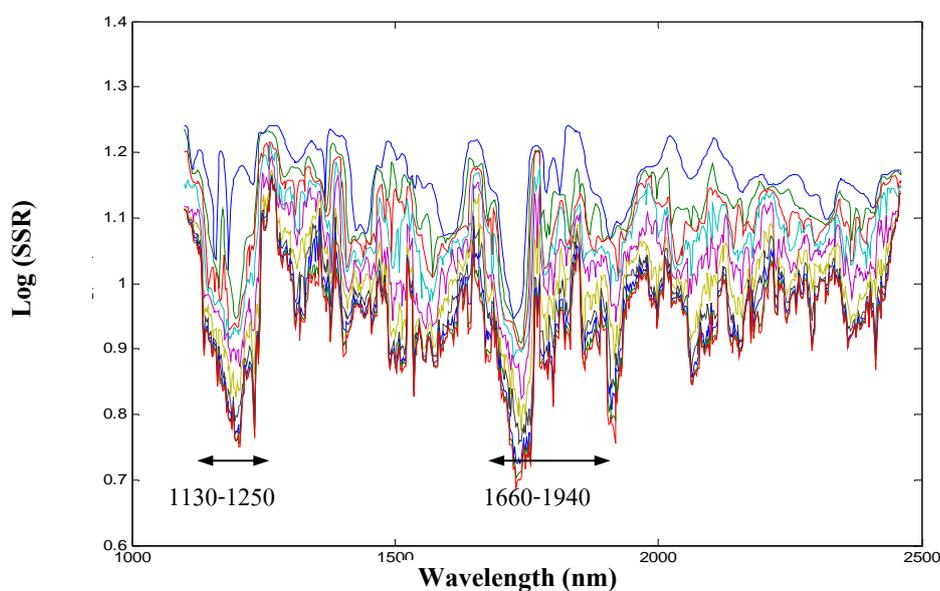


Figure 103 The residue lines obtained by MWPLSR for galactose

Two regions had low log (SSR) value which in the range of 1130-1250 nm and 1660-1940 nm. By using these regions and combination of both regions produced the calibration. The results were reported in Table 35.

Table 35 The statistical summaries of galactose content calibration model that vary wavelength regions.

| No | Region | Factor | Calibration | | Prediction | | RPD |
|----|-----------------------|--------|-------------|-------|------------|-------|-------|
| | | | R | SEC* | SEP* | Bias | |
| 1 | 1100-2500 | 1 | 0.809 | 0.310 | 0.201 | 0.016 | 1.990 |
| 2 | 1130-1250 | 2 | 0.701 | 0.376 | 0.266 | 0.036 | 1.504 |
| 3 | 1660-1940 | 3 | 0.751 | 0.240 | 0.245 | 0.011 | 1.633 |
| 4 | 1130-1250 + 1660-1940 | 2 | 0.709 | 0.372 | 0.260 | 0.030 | 1.538 |

* Unit (%W/W)

From the result, the calibration of the whole region provided the lowest SEP and the highest correlation coefficient. Therefore, the calibration from the whole regions was the best calibration as equation 46.

$$\% \text{ Galactose} = b_0 + \sum_{i=1100}^{2500} b_i x_i \quad (46)$$

Where b_0 is y-intercept of regression model, b is regression coefficient at wavelength i , X_i is the absorbance that was pretreated with MSC and second derivative, i is wavelength. The regression coefficient of this equation was shown in Figure 104.

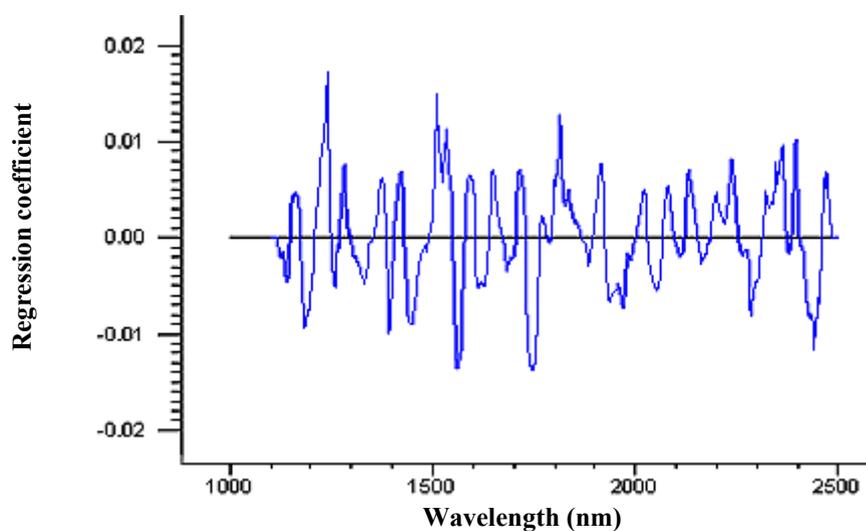


Figure 104 The regression coefficients of galactose from equation 46 with using wavelength in the region of 1100-2500 nm.

The correlation plots between the predicted value and actual of calibration set and validation set were presented in Figure 105.

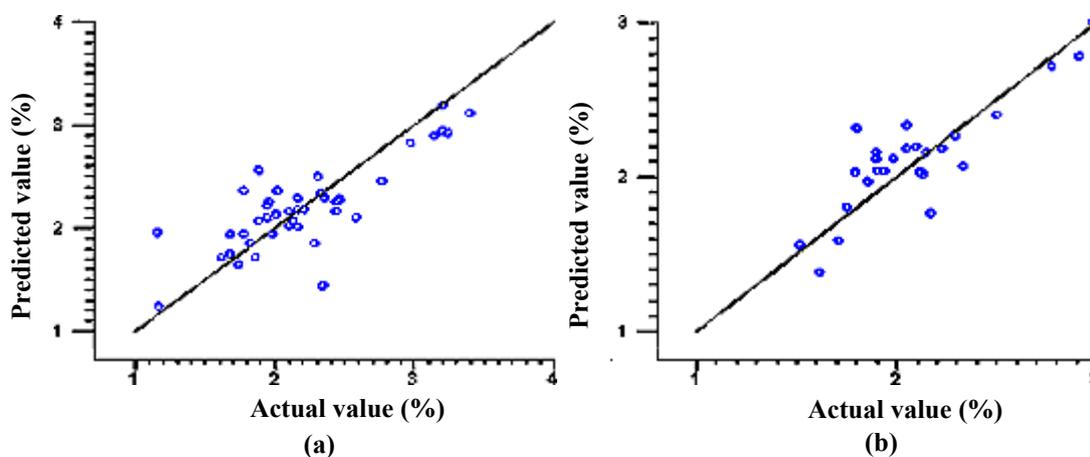


Figure 105 The correlation plots of galactose content between predicted value and actual value of calibration set (a) and validation set (b).

4.14 Mannose Content

4.14.1 Multiple Linear Regression (MLR) Analysis of Spectral Data and Mannose Content

The correlation of absorbance in the wavelength range at 110-2500 nm and mannose content in each wavelength was pretreated with second derivative to extract important information. The multiple correlation was shown in Figure 106.

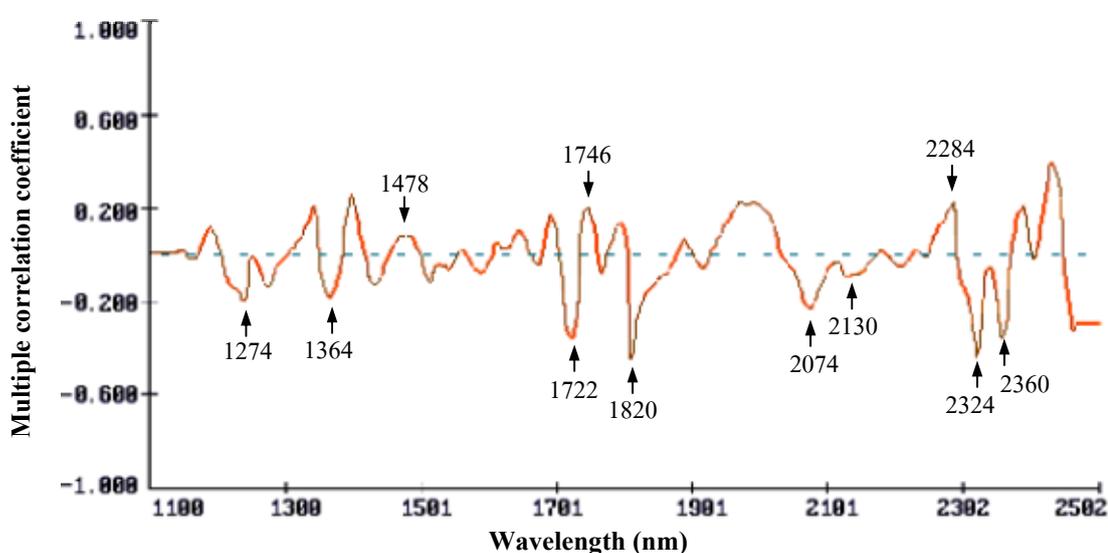


Figure 106 The multiple correlation coefficients of mannose content that pretreated with second derivative.

The first absorbance selection was selected from the correlations correspond near the second derivative band of mannose standard. The automatic selection of first absorbance was 1820 nm and the manual selection absorbance were 1274, 1364, 1478, 1722, 1746, 2074, 2130, 2284, 2324, and 2360 nm. The calibration model that obtained the lowest result of SEP, consisted of the first absorbance at 2324 nm. The followed absorbances selected by automatic selection were 1810, 1536, and 2358 nm, respectively. The tentative band at 2324 nm relate to C-H stretching combine CH_2 deformation. The band at 1820 nm relate to O-H stretching and 2nd overtone of C-O stretching combination. The band at 1540 nm corresponds to 1st overtone of O-H stretching. The tentative peaks at 2352 related to the 2nd of CH_2 bending (Burns, 2001). The correlation coefficient and standard error of calibration set were 0.805 and 0.129. The standard error of prediction and RPD were 0.104 and 1.279. The calibration for mannose contents analysis from NSAS program was obtained as equation 47.

$$\begin{aligned} \% \text{ Mannose} = & 1.125 - 114.557 \text{ d}^2\log(1/R_{2324}) - 1757.169 \text{ d}^2\log(1/R_{1820}) \\ & - 615.005 \text{ d}^2\log(1/R_{1536}) - 110.006 \text{ d}^2\log(1/R_{2358}) \end{aligned} \quad (47)$$

The correlation plots between the predicted value and actual value of calibration set and validation set showed in Figure 107.

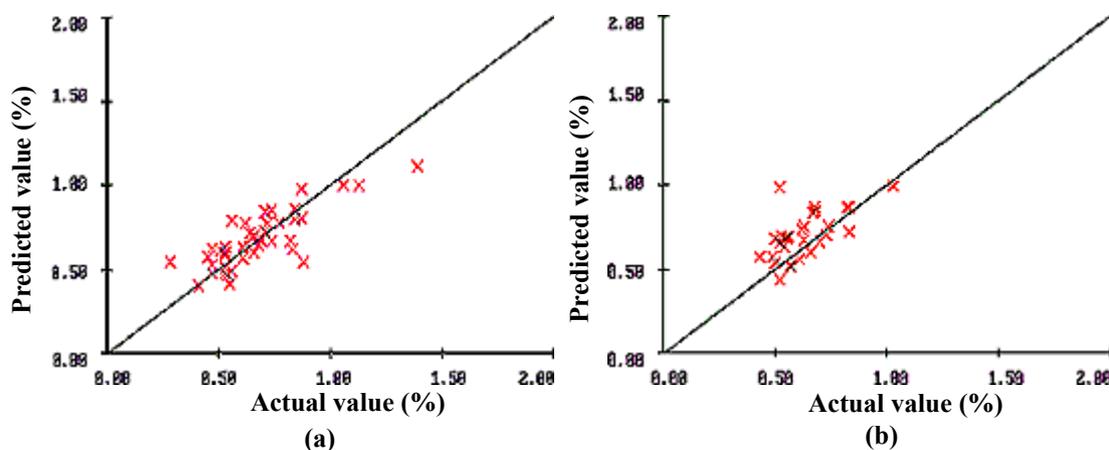


Figure 107 The correlation plots of mannose content between predicted value and actual value of calibration set (a) and validation set (b).

4.14.2 Partial Least Square Regression (PLSR) Analysis of Spectral Data and Mannose Content

This analysis was compared between different pretreatment before calibration performing for mannose content. The whole wavelengths in the region of 1100-2500 nm were calculated and validated as test set method. The statistical summaries were shown in Table 36.

Table 36 The statistical summaries of mannose content calibration model with pretreatment in vary method.

| No. | Pretreatment | Factor | Calibration | | Prediction | | RPD |
|-----|------------------------|--------|-------------|-------|------------|-------|-------|
| | | | R | SEC* | SEP* | Bias | |
| 1 | Average spectra | 3 | 0.248 | 0.237 | 0.137 | 0.041 | 0.949 |
| 2 | MSC | 2 | 0.702 | 0.147 | 0.121 | 0.046 | 1.074 |
| 3 | Second derivatives | 1 | 0.344 | 0.194 | 0.141 | 0.037 | 0.922 |
| 4 | MSC+ second derivative | 2 | 0.741 | 0.138 | 0.120 | 0.056 | 1.083 |

* Unit (%W/W)

The MSC and second derivative pretreatment spectra were calculated and evaluated. The results of this model showed the lowest values of SEP and SEC value. This pretreatment was selected, although that provided the highest

bias. The second derivative pretreated spectra gave higher the correlation coefficient. Therefore, the MSC and second derivative pretreatment spectra were selected to made calibration for mannose content. The spectra were transferred to moving window program. The result between log (SSR) and wavelength was illustrated in Figure 108.

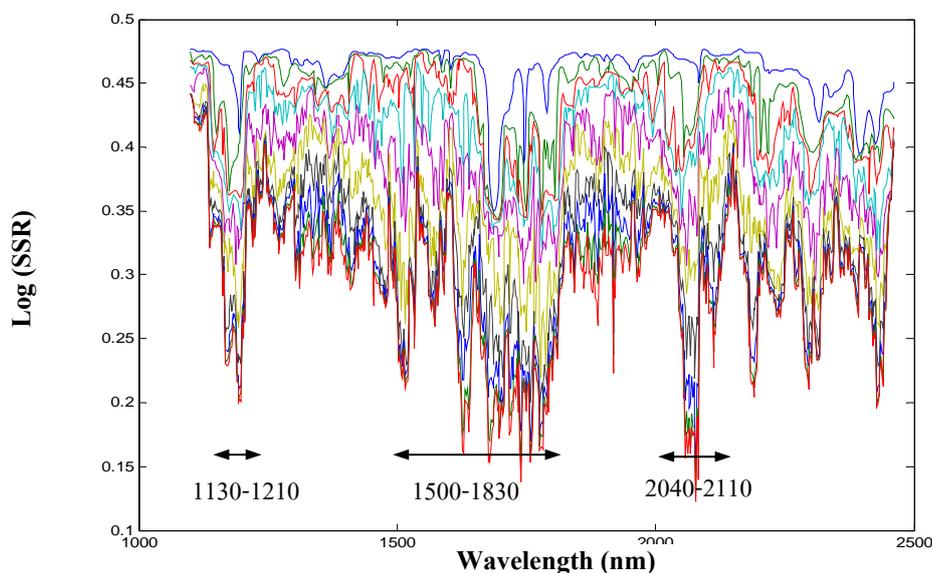


Figure 108 The residue lines obtained by MWPLSR for mannose content.

There were three regions which have low (SSR) value that in the region of 1130-1210 nm, 1500-1830 nm, and 2040-2380 nm. These regions and combination were used to produce the calibration. The statistic summaries presented in Table 37.

Table 37 The statistical summaries of mannose content calibration model that vary wavelength regions.

| No. | region | factor | Calibration | | Prediction | | RPD |
|-----|------------------------------------|--------|-------------|-------|------------|-------|-------|
| | | | R | SEC* | SEP* | Bias | |
| 1 | 1100-2500 | 2 | 0.741 | 0.138 | 0.120 | 0.056 | 1.083 |
| 2 | 1500-1830 | 1 | 0.839 | 0.112 | 0.116 | 0.068 | 1.121 |
| 3 | 2040-2110 | 7 | 0.657 | 0.156 | 0.117 | 0.079 | 1.111 |
| 4 | 1130-1210+1500-1830 | 1 | 0.648 | 0.156 | 0.110 | 0.064 | 1.182 |
| 5 | 1500-1830+2040-2110 | 7 | 0.756 | 0.135 | 0.116 | 0.031 | 1.130 |
| 6** | 1130-1210 + 1500-1830+2040-2110 | 5 | 0.823 | 0.117 | 0.108 | 0.067 | 1.204 |

* Unit (%W/W), ** Selected Calibration

The combination of three wavelength regions performed the best calibration. The correlation coefficient and SEC were 0.823, and 0.117. The SEP and RPD were 0.108 and 1.204 respectively. The best calibration was obtained as equation 48.

$$\% \text{ Mannose} = b_0 + \sum_{i=1130}^{1210} b_i x_i + \sum_{i=1500}^{1830} b_i x_i + \sum_{i=2040}^{2110} b_i x_i \quad (48)$$

b_0 is y-intercept of regression model, b is regression coefficient at wavelength i , x is $d^2 \log(1/R)$, i is wavelength. The regression coefficients of this equation were shown in Figure 109.

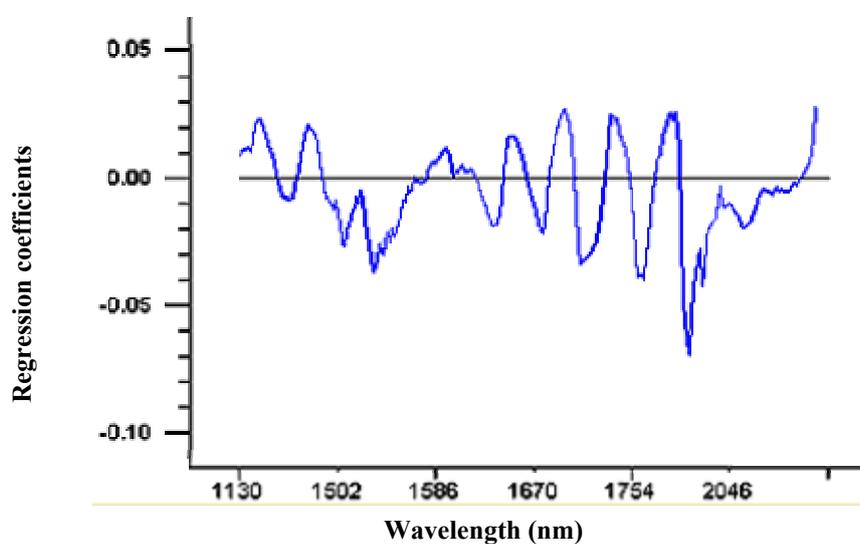


Figure 109 The correlation coefficients of mannose from equation 48 that used the combination of wavelength 1130-1210 nm, 1500-1830 nm, and 2040-2140 nm.

The correlation plots between the predicted value and actual value of calibration set and validation set presented in Figure 110.

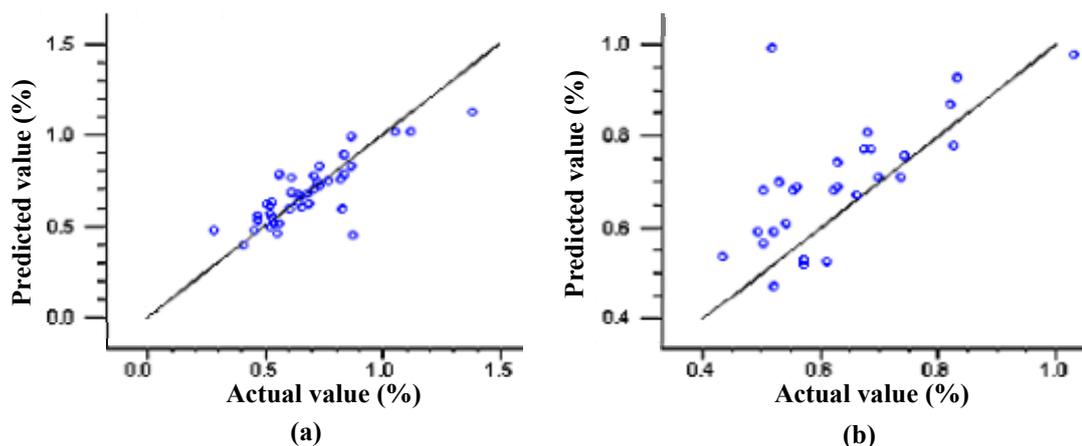


Figure 110 The correlation plots of mannose content between predicted value and actual value of calibration set(a) and validation set(b).

4.15 Comparisons of MLR and PLS Model

For the calibration performing in NIR spectroscopic method, most components were correlated with the NIR band excluding ash content. The ash content was a conventional expression for the residue on ignition from a solid sample according to certain testing procedures. As a result an ash number will indicate a quantitative measure for the quantity of mineral salts and other inorganic substances in the sample because it will be subject to not exactly defined thermal decomposition and volatility. The calibration for ash content was indirect correlation to amount of cellulose. The ash content from papers can be correlated with their NIR spectra, and the selected absorbance for calibration model that relate to O-H vibration (Anonymous. 2006).

MLR and PLS were developed to estimate the parameters of Eucalyptus wood powder. Each of the reference value give the optimum model in both of MLR and PLS model. The statistical results of correlation coefficients, SEC, and SEP were used for comparison and showed in Table 37.

The MLR calibration of total pulp yield, screen pulp yield, α -cellulose value, glucose, xylose, and arabinose content provided the better result than PLS calibration. For lignin, galactose, and mannose content showed the similar results by the PLS and MLR. For holocellulose, pentosan, 1% NaOH solubility, wood extractive, and ash results of PLS calibration presented better than MLR calibration.

In the PLS method with vary the different pretreatment, most the spectra were pretreated by second derivative spectra except lignin, wood extractives, arabinose, galactose, and mannose content.

The RPD for NIR calibration was the statistic for the parameter evaluation. It demonstrated how well the calibration models correlate with the chemical data. The RPD can classify the suitability of the chemical methods for NIR calibration according to the relationship between the error in analysis and the spread

in composition (Murray, 1986). The value was more over than three indicated the good calibration for prediction. In these experiment total pulp yield, screen pulp yield, and 1% NaOH solubility were higher RPD value.

Table 38 The summaries of statistic results in all components

| Property | Statistic | Pretreatment | Region | Factor | R | SEC* | SEP* | Bias | RPD |
|---------------------|-----------|--------------------------------|-----------------------------------|--------|-------|-------|-------|--------|-------|
| Total pulp yield | MLR | 2 nd derivative | 1820, 2368, 1238, 2152 | - | 0.964 | 0.833 | 0.845 | -0.144 | 3.178 |
| | PLS | 2 nd derivative | 1100-1250, 1372-1860 2030-2390 | 2 | 0.935 | 1.057 | 0.949 | 0.066 | 2.950 |
| Screen pulp yield | MLR | 2 nd derivative | 1820, 2368, 1238, 2152 | - | 0.967 | 0.804 | 0.843 | -0.178 | 3.357 |
| | PLS | 2 nd derivative | 1100-1250, 2030-2400 | 3 | 0.963 | 0.799 | 0.938 | -0.078 | 3.017 |
| Holocellulose | MLR | 2 nd derivative | 2338, 1474, 2176, 2454 | - | 0.693 | 1.690 | 1.650 | 0.519 | 1.261 |
| | PLS | 2 nd derivative | 2180-2300 | 7 | 0.779 | 1.400 | 1.496 | 0.245 | 1.391 |
| α -Cellulose | MLR | 2 nd derivative | 2276,1688,1396 | - | 0.869 | 1.220 | 0.677 | 0.031 | 2.496 |
| | PLS | 2 nd derivative | 1100-2500 | 5 | 0.913 | 0.962 | 0.772 | -0.103 | 2.189 |
| Pentosan | MLR | 2 nd derivative | 2240, 2320, 1604 | - | 0.773 | 0.888 | 0.761 | 0.240 | 1.367 |
| | PLS | 2 nd derivative | 1100-1830, 2190-2350 | 7 | 0.772 | 0.836 | 0.718 | 0.359 | 1.448 |
| Lignin | MLR | 2 nd derivative | 1672, 2430 | - | 0.833 | 0.746 | 0.442 | 0.275 | 1.697 |
| | PLS | MSC+2 nd derivative | 1480-1740, 2030-2500 | 2 | 0.807 | 0.781 | 0.426 | 0.325 | 1.761 |

Table 38 (Cont'd)

| Property | Statistic | Pretreatment | Region | Factor | R | SEC* | SEP* | Bias | RPD |
|----------------------|-----------|-------------------------------------|------------------------|--------|-------|-------|-------|--------|-------|
| 1%NaOH solubility | MLR | 2 nd derivative | 1934, 2186, 1888, 2316 | - | 0.945 | 1.090 | 0.702 | 0.044 | 3.704 |
| | PLS | 2 nd derivative | 1620-1820, 2150-2400 | 2 | 0.970 | 0.776 | 0.591 | 0.024 | 4.399 |
| Extractive | MLR | 2 nd derivative | 1762, 2132, 1708, 1612 | - | 0.925 | 0.384 | 0.436 | 0.288 | 2.094 |
| | PLS | MSC + 2 nd derivative | 1100-1970, 2090-2420 | 7 | 0.967 | 0.246 | 0.378 | 0.175 | 2.415 |
| Ash | MLR | 2 nd derivative | 1946, 1296, 1974, 1902 | - | 0.943 | 0.077 | 0.063 | 0.042 | 1.905 |
| | PLS | 2 nd derivative | 1100-1400, 1880-2080 | 7 | 0.965 | 0.058 | 0.061 | 0.027 | 1.967 |
| Glucose | MLR | 2 nd derivative | 2278, 1730, 2310, 1814 | - | 0.874 | 1.020 | 0.827 | -0.119 | 1.753 |
| | PLS | 2 nd derivative | 1430-1840, 2030-2400 | 2 | 0.800 | 1.195 | 0.867 | 0.106 | 1.672 |
| Xylose | MLR | 2 nd derivative | 1762, 2084, 1144 | - | 0.736 | 1.660 | 1.210 | -0.187 | 1.802 |
| | PLS | 2 nd derivative | 1200-1280, 1740-2090 | 2 | 0.720 | 1.648 | 1.280 | 0.205 | 1.703 |
| Arabinose | MLR | 2 nd derivative | 2266, 1240, 1152 | - | 0.843 | 0.095 | 0.086 | 0.019 | 1.628 |
| | PLS | MSC + 2 nd derivative | 1100-2500 | 2 | 0.783 | 0.106 | 0.092 | 0.012 | 1.522 |

Table 38 (Cont'd)

| Property | Statistic | Pretreatment | Region | Factor | R | SEC* | SEP* | Bias | RPD |
|-----------|-----------|--------------------------------|-----------------------------------|--------|-------|-------|-------|-------|-------|
| Galactose | MLR | 2 nd derivative | 2144, 1754, 2200 | - | 0.793 | 0.334 | 0.206 | 0.012 | 1.942 |
| | PLS | MSC+2 nd derivative | 1100-2500 | 1 | 0.809 | 0.310 | 0.201 | 0.016 | 1.990 |
| Mannose | MLR | 2 nd derivative | 2324, 1810, 1536, 2358 | - | 0.805 | 0.129 | 0.104 | 0.068 | 1.279 |
| | PLS | MSC+2 nd derivative | 1130-1210,1500-1830, 2040-2140 | 5 | 0.823 | 0.117 | 0.108 | 0.067 | 1.204 |

* Unit (%W/W)

CONCLUSIONS

Analysis of the pulp yield properties and the chemical compositions in *Eucalyptus camaldulensis* samples by the reflectance Near Infrared Spectroscopy (NIRs) were to develop the model that correlates between spectra and reference value. The PLS and MLR statistic were used for the calibration model development.

For the percentage of pulp yield properties, that could be divided into two values as total pulp yield, and screen pulp yield. MLR calibration model of both values provided R values of 0.964, and 0.967, respectively. The SEP values were 0.845, and 0.843, respectively. The statistic results of PLSR models presented R values of 0.935, and 0.963, respectively. The SEP values of these models were 0.949, and 0.938, respectively. In these properties, both of the MLR calibration model and the PLSR calibration model showed high correlation. However, the MLR model was more optimize than the PLSR calibration model.

The chemical compositions were carried out as holocellulose, α -cellulose, pentosan, lignin, 1% NaOH solubility, wood extractives, and ash content. The R value of MLR calibration models were 0.693, 0.869, 0.773, 0.833, 0.945, 0.925, and 0.943, respectively. The SEP of these statistic were 1.650, 0.677, 0.761, 0.442, 0.702, 0.436, and 0.063, respectively. PLSR calibration models provided the R value of these contents to 0.779, 0.913, 0.772, 0.803, 0.970, 0.967, and 0.965, respectively. It showed that both of MLR, and PLSR models for holocellulose, and pentosan prediction presented low correlations. The comparison between MLR, and PLSR model; MLR model for α -cellulose reported more precise than PLSR model. For holocellulose, pentosan, 1% NaOH solubility, wood extractives, and ash content indicated PLS calibrations were better. The others contents as lignin, galactose, and mannose showed the similar statistic results of MLR and PLSR.

Monosaccharide values were glucose, xylose, arabinose, galactose, and mannose. The spectra were calculated with MLR statistic that performed R values of 0.874, 0.736, 0.843, 0.793, and 0.805, respectively. SEP values of these models were 0.827, 1.212, 0.086, 0.206, and 0.104. For PLSR models of these values, R values were 0.800, 0.720, 0.783, 0.809, and 0.823, respectively. SEP values of these models were 0.867, 1.280, 0.092, 0.201, and 0.108. The MLR calibration model of glucose, xylose, and arabinose contents were better than PLSR calibration. The statistical results of PLSR and MLR calibration were not different for prediction of galactose and mannose content.

In additional, RPD results over 3.0, in which confirms high RPD that NIR spectroscopy could accuracy predict total pulp yield, screen pulp yield, and 1% NaOH solubility. Moreover it was possible to examine α -cellulose, lignin, wood extractives, ash, glucose, arabinose, galactose, and mannose content by NIRs. Howere, the calibration of some properties showed that low RPD could be improved with more intensive research such as increasing number of samples in terms of varieties, location of plantation, and season of growing as well as the age of plant.

LITERATURE CITED

- Anonymous. 1998. **SESAME Software Operation manual** BRAN+LUEBBE. p.280-282
- _____. 2001. **InfraAlyzer 500 and SESAME-Software version 3.10, Training Manual.** BRAN+LUEBBE. p.128
- _____. 2006. Determination of Ash from Paper. **Bruker Optics.** Available Source: <http://www.brukeroptics.com/applications/paper.html>, February, 2, 2006.
- Alen, R. 2000. **Forest Products Chemistry.** edited by J. Gullichsen and H. Paulapuro. Fapet Oy. Finland.
- Ali, M., A.M. Emsley, H. Herman, and R.J. Heywood. 2001. Spectroscopic studies of the ageing of cellulosic paper. **Polymer.** 42 : p.2893-2900
- Burns, D.A. 2001. **Handbook of Near-Infrared Analysis: Historical Development.** Marcel Dekker, Inc. New York. p.814
- Biermann, C.J. 1996. **Handbook of pulping and papermaking.** Academic press, Inc., U.K. p. 86-91.
- Boysworth, M.K. and K.S. Booksh. 2001. Aspects of multivariate calibration applied to near-infrared spectroscopy. In **Handbook of Near-Infrared Analysis.** edited by D.A.Burns and E.W.Ciurczak, 2nd edition. Marcel Dekker, Inc. New York. p.212.
- Browning, B.L. **Methods of Wood Chemistry.** Vol.2. Wiley Interscience. New York. 1967. Chap.19
- Chaikumpollert, O. 2004. **Isolation and Structural Characterisation of Hemicelluloses from Vetiver Grass.** M.S.. Thesis, Mahidol University. Bangkok. p.56.
- Cao, B., U.Tshirner, S. Ramaswamy and A.Webb. 1997. A rapid modified gas chromatographic method for carbohydrate analysis of wood pulps. **Tappi J.** vol.80(9) : 193-197.
- Chalmers, J.M. 2000. Spectroscopy in process analysis. Sheffield Academic Press. Cleveland. UK. p.53-56.
- Connors, T.E., and S. Banerijee. 1995. **Surface analysis of paper.** CRC Press. U.S.A. p.198.
- Easty, D.B., S.A. Berben, F.A. DeThomas, P.J. Brimmer. 1990. Near infrared spectroscopy for the analysis of wood pulp: Quantifying hardwood-softwood mixtures and estimating lignin content. **Tappi J.** (10): 257-261.

- Eldridge K., J. Davidson, C. Harwood and G.V. WYK 1997. **Eucalyptus Domestication And Breeding**, Clarendon Press. Oxford. p.11-20.
- Hruschka, W.R. 1990. Data Analysis: Wavelength Selection Methods. *In* P. Williams and K. Norris., eds. *Near-Infrared Technology in the Agricultural and Food Industries*. American Association of Cereal Chemists, Inc. USA. p.47.
- Kasemsumran, S. 2005. **Nondestructive Quantitative Analysis of Biomedical Sample by Near-Infrared Spectroscopy with Chemometrics and Development of Novel Chemometric Method for Quantitative Analysis**. PH.D. Thesis, Kwansai Gakuin University. Japan. p.13-31.
- Kawano, S. 2001. Reserch Activity of Nondestructive Quality Evaluation Techniques and NIR Spectroscopy in Japan. *In* Warunee Thanapase, Anupun Terdwongworakul and Thongchai Suwansichon. การอบรมเชิงปฏิบัติการการควบคุมคุณภาพสินค้าด้วยเทคนิค **Near Infrared Spectroscopy** เพื่อการแข่งขันในเวทีการค้าโลก. Edited 1. Kasetsart Agricultural and Agro-Industrail Product Improvement Institute, Kasetsart University, Bangkok. p.17-23.
- Kelley, S.S., T.G. Rials, R. Snell, L.H. Groom, and A. Sluiter. 2004. Use of Near Infrared spectroscopy to mesure the chemical and mechanicl properties of solid wood. **Wood Sci Technol**. 38:257-276.
- Marklund, A., J.B. Hauksson, U. Edlund and M. Sjostrom. 1999. Multivariate data analysis base on orthogonal signal correction and near infrared spectroscopy. **Nordic Pulp & Paper Research Journal**. 14(2). p. 140-148.
- Meghanathan, N., V.M. Saucedo and G.A. Krishnagopalan. 2000. Modeling of kraft pulping process based on on-line data from near infrared analyzer. AIChE Symposium Series 324. **Fundamentals and Numerical Modeling of Unit Operations in the Forest Products Industries**. p. 36-41.
- Michell, A.J. 1995. Pulpwood Quality Estimation by Near Infrared Spectroscopic Measurements on Eucalypt Woods. **Apitta J**. Vol 48(6):425-428
- _____ and L.R.Schimleck 1996. NIR Spectroscopy of woods from Eucalyptus globules. **Apitta J**. Vol 49(1):23-26.
- Murray, I. 1986 The NIR spectra of homologous series of organic compounds. *In*: J. Hollo, K.J. Kaffka and J.L. Gonczy, Editors, **Proceedings international NIR conference**. Akademiai Kiado. Budapest. pp. 13-28
- _____ 1990. Application of NIRS in Agriculture. *In* M. Iwamoto and S. Kawano. eds. **The Proceedings of The Second International Near Infrared Spectroscopy Conference**. Korin Publishing Co., Ltd. p.11-20.

- Norris, K. and W. Hruschka 1982. Least-squares curve fitting of near-infrared spectra predicts protein and moisture content of ground wheat. **Appl. Spectrosc.** 36. p. 261-265.
- Osborne, B.G., T. Fearn and P.H. Hindle 1993. **Practical NIR Spectroscopy with Application in Food and Beverage Analysis**. 2nd Edition. Long Scientific and Technical, Singapore. p. 29-38
- Ozaki, Y. 2003. Introduction to NIR Spectroscopy-Principles, Instrumentation and Application. *In* Warunee Thanapase, Anupun Terdwongworakul and Thongchai Suwansichon.การอบรมเชิงปฏิบัติการการควบคุมคุณภาพสินค้าด้วยเทคนิค**Near Infrared Spectroscopy** เพื่อการแข่งขันในเวทีการค้าโลก. Edited 1. Kasetsart Agricultural and Agro-Industral Product Improvement Institute, Kasetsart University, Bangkok. p. 17.
- _____, S. Sasic and J.H. Jiang.2001. How we unravel complicated near infrared spectra?-Recent progress in spectra analysis methods for resolution enhancement and band assignments in the near infrared region. *J.Near-Infrared Spectroscopy*. 9:63-65.
- Reiter, C., B. Reinhardt, B. Borchers, and P. Patrick. 1999. Characterization of pulp and chemical additives with the help of FT-NIR technique. *Papier* (Bingen, Germany) 53 (10A). p. V20-V24.
- Rodrigues, T., J.O. Faix and H. Pereira 2001. Determination of Monosaccharide Composition of *Eucalyptus globules* wood by FTIR Spectroscopy. **Holzforchung** 55. p. 263-269.
- Schimleck, L.R and A. Michell, A.J. 1998. Determintion of within-Tree Variation of krft pulp yield using Near-Infrared spectroscopy. **Tappi J.** Vol.81(9). p. 229-236
- _____, P.J.Wright, A. Michell. J.Wallis, and F.A. Adrian. 1997. Near-Infrared spectra and chemical compositions of *E. globulus* and *E. nitens* plantation woods. **Appita Journal**. Vol.50(1). p. 40-46.
- _____ and R. Evans. 2004. Estimation of Pinus radiate D.Don trachied morphological characteristics by Near Infrared spectroscopy. **Holzforchung** 58. p. 66-73.
- _____ and Y. Yazaki 2003. Analysis of *Pinus radiata* D. Don Bark by Near Infrared Spectroscopy . **Holzforchung** 57. p. 520-526.
- Schultz, T.P., and D.A. Burns. 1990. Rapid secondary analysis of lignocelluloses: comparison of near infrared(NIR) and Fourier transform infrared(FTIR). **Tappi J.** 5. p. 209-212
- Shenk J.S., J.J. Workman, Jr and M.O. Westerhaus. 1992. **Handbook of Near-Infrared Analysis**.Marcel Dekker, Inc. New York. p. 383-431.

- Shimoyama, M., S. Hayano, K. Matsukawa, H. Inoue, T. Ninomiya, and Y. Ozaki. 1998. Discrimination of ethylene/vinyl acetate copolymer with different composition and prediction of content of vinyl acetate in copolymers and their melting points by near infrared spectroscopy and chemometrics. **J Polym. Sci. Part B: Polym. Phys.** 36. p. 1529-1537
- Siesler, H.W., Y. Ozaki, S. Kawata 2002. **Near-Infrared Spectroscopy. Principles, Instruments, Applications.** WILEY-VCH. Germany. p.208.
- Sjostrom, E. 1993. **Wood Chemistry Fundamentals and Applications.** 2nd Edition. Academic press, Inc. London. p.105-130.
- Suwonsichon, T. Data Analysis of NIR Spectra: Quality and Qualitative Analysis. In Warunee Thanapase, Anupun Terdwongworakul and Thongchai Suwonsichon. การอบรมเชิงปฏิบัติการการควบคุมคุณภาพสินค้าด้วยเทคนิค **Near Infrared Spectroscopy** เพื่อการแข่งขันในเวทีการค้าโลก. Edited I. Kasetsart Agricultural and Agro-Industrial Product Improvement Institute, Kasetsart University, Bangkok. p.99.
- TAPPI (Technical Association of the Pulp and Paper Industry), 1983a. **Acid-insoluble lignin in wood and pulp.** Official Test Method T-222 om-8. Tappi press. Atlanta.
- TAPPI (Technical Association of the Pulp and Paper Industry), 1983a. **Ash in wood, pulp, paper and paperboard: combustion at 525°C.** Official Test Method T-211 om-93. Tappi press. Atlanta.
- TAPPI (Technical Association of the Pulp and Paper Industry), 1983a. **Alpha-, Beta-, and Gamma-Cellulose in Pulp.** Official Test Method T-203 om-93. Tappi press. Atlanta.
- TAPPI (Technical Association of the Pulp and Paper Industry), 1983a. **Analysis of pentosan content.** Official Test Method T 223 cm-84. Tappi press. Atlanta.
- TAPPI (Technical Association of the Pulp and Paper Industry), 1983a. **One percent sodium hydroxide solubility of wood and pulp.** Official Test Method T 223 cm-84 Tappi press. Atlanta.
- Terdwongworakul. A. 2002. Pre-treatment of spectra. การอบรมเชิงปฏิบัติการการควบคุมคุณภาพสินค้าด้วยเทคนิค **Near Infrared Spectroscopy** เพื่อการแข่งขันในเวทีการค้าโลก. Edition 1., Kasetsart agricultural and Agro-industrial Product Improvement Institute, Bangkok. p.67-98.
- Thomas, E.V. 1994. A primer on multivariate calibration. **Anal.Chem.** 66: 795A-804A.
- Tsuchikawa, S., K. Inoue, J. Noma and K. Hyshi. 2003. Application of Near-Infrared Spectroscopy to Wood Discrimination. **J. Wood Sci.** 49 p. 29-35

- Wallbacks, L., U. Edlund, and B. Norden. 1991. Multivariate Characterization of Pulp Using solid-state ^{13}C NMR, FTIR, and NIR **Tappi J.** (10). p. 201- 206
- Wenhao, S., W.Xinsheng, and L. Huanbin. 2004. Study om-line pulp Kappa number determination with spectroscopy during batch pulping. Appita Annual Conference and Exhibition, 58th, (vol. 2). p.50-505.
- Wright, J.A. 1990. Prediction of pulp yield and cellulose content from wood samples using near infrared reflectance spectroscopy. **Tappi J.** (8). p.164-166
- Williams, P.C., and K. Norris 2001. **Near Infrared Technology in the Agricultural and Food Industries.** 2nd Edition. American Association of Cereal. p.164-165.
- Xiaomei L., W, Shuangfei. 2003. Application of near-infrared spectroscopy in pulp and paper industry. Zhongguo Zaozhi Xuebao. 18 (2). p. 189-1944

APPENDIX

APPENDIX A

Appendix Table A1 The statistic result of total pulp yield calibration.

| No. | Wavelength (nm) | | | | R | SEC* | SEP* | Bias | RPD |
|-----|-------------------|-------------------|-------------------|-------------------|--------|-------|-------|--------|-------|
| 1 | 1778 ^M | | | | -0.648 | 2.300 | 2.260 | 0.249 | 1.239 |
| 2 | 1778 ^M | 2368 ^C | | | 0.943 | 1.020 | 1.020 | -0.058 | 2.745 |
| 3 | 1778 ^M | 2368 ^C | 2294 ^C | | 0.952 | 0.952 | 0.913 | -0.006 | 3.067 |
| 4 | 1778 ^M | 2368 ^C | 2294 ^C | 1242 ^C | 0.959 | 0.892 | 0.895 | -0.081 | 3.128 |
| 5 | 1820 ^M | | | | -0.378 | 2.790 | 2.430 | 0.038 | 1.152 |
| 6 | 1820 ^M | 2368 ^C | | | 0.939 | 1.050 | 1.010 | -0.021 | 2.772 |
| 7 | 1820 ^M | 2368 ^C | 1238 ^C | | 0.952 | 0.944 | 0.923 | -0.060 | 3.034 |
| 8** | 1820 ^M | 2368 ^C | 1238 ^C | 2152 ^C | 0.964 | 0.833 | 0.845 | -0.154 | 3.314 |
| 9 | 2100 ^M | | | | -0.785 | 1.870 | 1.600 | -0.234 | 1.750 |
| 10 | 2100 ^M | 2370 ^C | | | 0.940 | 1.040 | 0.963 | 0.076 | 2.908 |
| 11 | 2100 ^M | 2370 ^C | 1240 ^C | | 0.954 | 0.931 | 0.918 | -0.033 | 3.050 |
| 12 | 2100 ^M | 2370 ^C | 1240 ^C | 2072 ^C | 0.964 | 0.833 | 0.961 | -0.009 | 2.914 |
| 13 | 2266 ^M | | | | -0.736 | 2.040 | 1.620 | -0.207 | 1.728 |
| 14 | 2266 ^M | 2368 ^C | | | 0.945 | 1.000 | 0.930 | -0.053 | 3.011 |
| 15 | 2266 ^M | 2368 ^C | 1240 ^C | | 0.954 | 0.928 | 0.881 | -0.144 | 3.178 |
| 16 | 2266 ^M | 2368 ^C | 1240 ^C | 1988 ^C | 0.965 | 0.825 | 0.964 | -0.123 | 2.905 |
| 17 | 2332 ^M | | | | -0.784 | 1.870 | 1.620 | 0.141 | 1.728 |
| 18 | 2332 ^M | 2368 ^C | | | 0.945 | 1.000 | 0.945 | -0.059 | 2.963 |
| 19 | 2332 ^M | 2368 ^C | 2018 ^C | | 0.953 | 0.941 | 0.982 | -0.057 | 2.851 |
| 20 | 2332 ^M | 2368 ^C | 2018 ^C | 1530 ^C | 0.959 | 0.887 | 1.040 | -0.037 | 2.692 |
| 21 | 2368 ^C | | | | -0.938 | 1.050 | 0.964 | -0.067 | 2.905 |
| 22 | 2368 ^C | 2342 ^C | | | 0.948 | 0.973 | 0.980 | -0.075 | 2.857 |
| 23 | 2368 ^C | 2342 ^C | 1398 ^C | | 0.956 | 0.908 | 1.020 | 0.108 | 2.745 |
| 24 | 2368 ^C | 2342 ^C | 1398 ^C | 1890 ^C | 0.958 | 0.902 | 0.962 | -0.091 | 2.911 |

Appendix Table A2 The statistic result of the screen pulp yield calibration.

| No. | Wavelength (nm) | | | | R | SEC* | SEP* | Bias | RPD |
|-----|-------------------|-------------------|-------------------|-------------------|--------|-------|-------|--------|-------|
| 1 | 1778 ^M | | | | -0.658 | 2.270 | 2.333 | 0.344 | 1.213 |
| 2 | 1778 ^M | 2368 ^C | | | 0.946 | 0.993 | 1.020 | -0.053 | 2.775 |
| 3 | 1778 ^M | 2368 ^C | 2296 ^C | | 0.955 | 0.921 | 0.907 | -0.033 | 3.120 |
| 4 | 1778 ^M | 2368 ^C | 2296 ^C | 1242 ^C | 0.963 | 0.840 | 0.870 | -0.166 | 3.253 |
| 5 | 1820 ^M | | | | -0.383 | 2.780 | 2.440 | 0.009 | 1.160 |
| 6 | 1820 ^M | 2368 ^C | | | 0.942 | 1.020 | 1.010 | -0.038 | 2.802 |
| 7 | 1820 ^M | 2368 ^C | 1238 ^C | | 0.956 | 0.911 | 0.918 | -0.116 | 3.083 |
| 8** | 1820 ^M | 2368 ^C | 1238 ^C | 2152 ^C | 0.967 | 0.804 | 0.843 | -0.178 | 3.357 |
| 9 | 2100 ^M | | | | -0.791 | 1.840 | 1.610 | -0.216 | 1.758 |
| 10 | 2100 ^M | 2370 ^C | | | 0.944 | 1.010 | 0.965 | 0.071 | 2.933 |
| 11 | 2100 ^M | 2370 ^C | 1240 ^C | | 0.959 | 0.877 | 0.929 | -0.099 | 3.046 |
| 12 | 2100 ^M | 2370 ^C | 1240 ^C | 2072 ^C | 0.968 | 0.793 | 0.964 | -0.098 | 2.936 |
| 13 | 2266 ^M | | | | -0.739 | 2.030 | 1.620 | 0.222 | 1.747 |
| 14 | 2266 ^M | 2368 ^C | | | 0.949 | 0.965 | 0.932 | -0.061 | 3.036 |
| 15 | 2266 ^M | 2368 ^C | 1240 ^C | | 0.959 | 0.877 | 0.885 | -0.205 | 3.198 |
| 16 | 2266 ^M | 2368 ^C | 1240 ^C | 1988 ^C | 0.968 | 0.784 | 0.963 | -0.200 | 2.939 |
| 17 | 2332 ^M | | | | -0.786 | 1.860 | 1.630 | 0.178 | 1.736 |
| 18 | 2332 ^M | 2368 ^C | | | 0.948 | 0.972 | 0.947 | -0.062 | 2.988 |
| 19 | 2332 ^M | 2368 ^C | 2018 ^C | | 0.955 | 0.918 | 0.951 | 0.034 | 2.976 |
| 20 | 2332 ^M | 2368 ^C | 2018 ^C | 1530 ^C | 0.962 | 0.854 | 0.917 | -0.068 | 3.086 |
| 21 | 2368 ^C | | | | -0.942 | 1.020 | 0.972 | -0.079 | 2.912 |
| 22 | 2368 ^C | 2296 ^C | | | 0.954 | 0.915 | 0.941 | -0.030 | 3.007 |
| 23 | 2368 ^C | 2296 ^C | 1242 ^C | | 0.963 | 0.833 | 0.903 | -0.015 | 3.134 |
| 24 | 2368 ^C | 2296 ^C | 1242 ^C | 1636 ^C | 0.968 | 0.786 | 0.846 | -0.281 | 3.345 |

Appendix Table A3 The statistic result of the holocellulose calibration

| No. | Wavelength (nm) | | | | R | SEC* | SEP* | Bias | RPD |
|-----|-------------------|-------------------|-------------------|-------------------|--------|-------|-------|-------|-------|
| 1 | 1162 ^M | | | | -0.587 | 1.830 | 2.000 | 0.652 | 1.040 |
| 2 | 1162 ^M | 1288 ^C | | | 0.670 | 1.700 | 1.850 | 0.531 | 1.124 |
| 3 | 1162 ^M | 1288 ^C | 1512 ^C | | 0.683 | 1.690 | 1.850 | 0.613 | 1.124 |
| 4 | 1162 ^M | 1288 ^C | 1512 ^C | 2454 ^C | 0.704 | 1.670 | 1.770 | 0.604 | 1.175 |
| 5 | 1220 ^M | | | | -0.520 | 1.930 | 1.970 | 0.320 | 1.056 |
| 6 | 1220 ^M | 1418 ^C | | | 0.688 | 1.660 | 1.920 | 0.707 | 1.083 |
| 7 | 1220 ^M | 1418 ^C | 2456 ^C | | 0.711 | 1.630 | 1.810 | 0.626 | 1.149 |
| 8 | 1220 ^M | 1418 ^C | 2456 ^C | 1794 ^C | 0.729 | 1.610 | 1.850 | 0.534 | 1.124 |
| 9 | 1288 ^M | | | | 0.670 | 1.680 | 1.850 | 0.534 | 1.124 |
| 10 | 1288 ^M | 1732 ^C | | | 0.683 | 1.670 | 1.840 | 0.570 | 1.130 |
| 11 | 1288 ^M | 1732 ^C | 1514 ^C | | 0.698 | 1.660 | 1.850 | 0.642 | 1.124 |
| 12 | 1288 ^M | 1732 ^C | 1514 ^C | 2452 ^C | 0.714 | 1.650 | 1.780 | 0.607 | 1.169 |
| 13 | 1346 ^M | | | | -0.340 | 2.150 | 2.080 | 0.352 | 1.000 |
| 14 | 1346 ^M | 1790 ^C | | | 0.672 | 1.700 | 1.820 | 0.530 | 1.143 |
| 15 | 1346 ^M | 1790 ^C | 2454 ^C | | 0.687 | 1.680 | 1.830 | 0.560 | 1.137 |
| 16 | 1346 ^M | 1790 ^C | 2454 ^C | 2294 ^C | 0.701 | 1.680 | 1.830 | 0.644 | 1.137 |
| 17 | 1424 ^M | | | | 0.640 | 1.740 | 1.990 | 0.933 | 1.045 |
| 18 | 1424 ^M | 1292 ^C | | | 0.678 | 1.680 | 1.880 | 0.668 | 1.106 |
| 19 | 1424 ^M | 1292 ^C | 2452 ^C | | 0.690 | 1.680 | 1.820 | 0.642 | 1.143 |
| 20 | 1424 ^M | 1292 ^C | 2452 ^C | 1798 ^C | 0.712 | 1.650 | 1.940 | 0.820 | 1.072 |
| 21 | 1482 ^M | | | | -0.591 | 1.820 | 1.760 | 0.330 | 1.182 |
| 22 | 1482 ^M | 1288 ^C | | | 0.673 | 1.690 | 1.820 | 0.500 | 1.143 |
| 23 | 1482 ^M | 1288 ^C | 1732 ^C | | 0.688 | 1.680 | 1.800 | 0.529 | 1.156 |
| 24 | 1482 ^M | 1288 ^C | 1732 ^C | 1512 ^C | 0.703 | 1.670 | 1.820 | 0.636 | 1.143 |

Appendix Table A3 (Cont'd)

| No. | Wavelength (nm) | | | R | SEC* | SEP* | Bias | RPD | |
|------|-------------------|-------------------|-------------------|-------------------|-------|-------|-------|-------|-------|
| 25 | 1598 ^M | | | -0.451 | 2.020 | 1.810 | 0.142 | 1.149 | |
| 26 | 1598 ^M | 1288 ^C | | 0.670 | 1.700 | 1.810 | 0.538 | 1.149 | |
| 27 | 1598 ^M | 2390 ^C | 1682 ^C | 0.684 | 1.690 | 1.890 | 0.674 | 1.101 | |
| 28 | 1598 ^M | 2390 ^C | 1682 ^C | 2340 ^C | 0.697 | 1.690 | 1.940 | 0.785 | 1.072 |
| 29 | 1794 ^M | | | -0.593 | 1.820 | 1.740 | 0.393 | 1.195 | |
| 30 | 1794 ^M | 1288 ^C | | 0.670 | 1.700 | 1.860 | 0.543 | 1.118 | |
| 31 | 1794 ^M | 1288 ^C | 1734 ^C | 0.689 | 1.680 | 1.920 | 0.651 | 1.083 | |
| 32 | 1794 ^M | 1288 ^C | 1734 ^C | 1516 ^C | 0.702 | 1.670 | 1.930 | 0.706 | 1.078 |
| 34 | 1918 ^M | | | -0.589 | 1.830 | 2.030 | 1.070 | 1.025 | |
| 35 | 1918 ^M | 1292 ^C | | 0.684 | 1.670 | 1.890 | 0.721 | 1.101 | |
| 36 | 1918 ^M | 1292 ^C | 1798 ^C | 0.696 | 1.660 | 1.960 | 0.856 | 1.061 | |
| 37 | 1918 ^M | 1292 ^C | 1798 ^C | 2456 ^C | 0.720 | 1.630 | 1.920 | 0.839 | 1.083 |
| 38 | 2122 ^M | | | -0.511 | 1.940 | 1.820 | 0.423 | 1.143 | |
| 39 | 2122 ^M | 1282 ^C | | 0.677 | 1.680 | 1.830 | 0.620 | 1.137 | |
| 40 | 2122 ^M | 1282 ^C | 1730 ^C | 0.694 | 1.670 | 1.780 | 0.594 | 1.169 | |
| 41 | 2122 ^M | 1282 ^C | 1730 ^C | 1796 ^C | 0.701 | 1.680 | 1.840 | 0.639 | 1.130 |
| 42 | 2274 ^C | | | -0.408 | 2.060 | 1.870 | 0.276 | 1.112 | |
| 43 | 2274 ^C | 1288 ^C | | 0.671 | 1.700 | 1.840 | 0.515 | 1.130 | |
| 44 | 2274 ^C | 1288 ^C | 1696 ^C | 0.694 | 1.670 | 1.840 | 0.662 | 1.130 | |
| 45 | 2274 ^C | 1288 ^C | 1696 ^C | 2156 ^C | 0.705 | 1.670 | 1.800 | 0.674 | 1.156 |
| 46 | 2338 ^M | | | -0.464 | 2.000 | 1.970 | 0.347 | 1.056 | |
| 47 | 2338 ^M | 1474 ^C | | 0.621 | 1.790 | 1.830 | 0.510 | 1.137 | |
| 48 | 2338 ^M | 1474 ^C | 2176 ^C | 0.668 | 1.730 | 1.700 | 0.600 | 1.224 | |
| 49** | 2338 ^M | 1474 ^C | 2176 ^C | 2454 ^C | 0.693 | 1.690 | 1.650 | 0.519 | 1.261 |

Appendix Table A4 The statistic result of the α -cellulose calibration

| No. | Wavelength (nm) | | | | R | SEC* | SEP* | Bias | RPD |
|------|-------------------|-------------------|-------------------|-------------------|--------|-------|-------|--------|-------|
| 1 | 1490 ^M | | | | -0.649 | 1.820 | 1.280 | -0.189 | 1.320 |
| 2 | 1490 ^M | 2316 ^C | | | 0.794 | 1.480 | 1.040 | -0.109 | 1.625 |
| 3 | 1490 ^M | 2316 ^C | 1732 ^C | | 0.842 | 1.330 | 0.939 | -0.169 | 1.800 |
| 4 | 1490 ^M | 2316 ^C | 1732 ^C | 1266 ^C | 0.875 | 1.210 | 0.795 | -0.494 | 2.126 |
| 5 | 1785 ^M | | | | -0.648 | 1.820 | 1.200 | 0.104 | 1.408 |
| 6 | 1785 ^M | 2316 ^C | | | 0.765 | 1.560 | 1.050 | 0.109 | 1.610 |
| 7 | 1785 ^M | 2316 ^C | 1768 ^C | | 0.876 | 1.340 | 0.953 | 0.065 | 1.773 |
| 8 | 1785 ^M | 2316 ^C | 1768 ^C | 1162 ^C | 0.864 | 1.250 | 0.910 | -0.021 | 1.857 |
| 9 | 1820 ^M | | | | 0.012 | 2.390 | 1.680 | 0.542 | 1.006 |
| 10 | 1820 ^M | 1728 ^C | | | 0.762 | 1.570 | 1.270 | 0.506 | 1.331 |
| 11 | 1820 ^M | 1728 ^C | 1686 ^C | | 0.876 | 1.190 | 0.862 | 0.194 | 1.961 |
| 12 | 1820 ^M | 1728 ^C | 1686 ^C | 1818 ^C | 0.890 | 1.140 | 0.930 | 0.240 | 1.817 |
| 13 | 2276 ^M | | | | -0.688 | 1.740 | 0.968 | 0.073 | 1.746 |
| 14 | 2276 ^M | 1688 ^M | | | 0.822 | 1.380 | 0.907 | -0.038 | 1.863 |
| 15** | 2276 ^M | 1688 ^M | 1396 ^M | | 0.869 | 1.220 | 0.677 | 0.031 | 2.496 |
| 16 | 2276 ^M | 1688 ^M | 1396 ^M | 2462 ^C | 0.891 | 1.130 | 0.754 | -0.005 | 2.241 |
| 17 | 2312 ^C | | | | 0.747 | 1.590 | 1.040 | 0.147 | 1.625 |
| 18 | 2312 ^C | 1686 ^C | | | 0.854 | 1.260 | 0.954 | 0.056 | 1.771 |
| 19 | 2312 ^C | 1686 ^C | 1544 ^C | | 0.877 | 1.180 | 0.855 | -0.008 | 1.977 |
| 20 | 2312 ^C | 1686 ^C | 1544 ^C | 1758 ^C | 0.883 | 1.170 | 0.851 | -0.006 | 1.986 |
| 21 | 2338 ^M | | | | -0.593 | 1.930 | 0.877 | 4.790 | 1.927 |
| 22 | 2338 ^M | 1728 ^C | | | 0.828 | 1.360 | 0.950 | 0.174 | 1.779 |
| 23 | 2338 ^M | 1728 ^C | 2280 ^C | | 0.855 | 1.270 | 0.864 | 1.141 | 1.956 |
| 24 | 2338 ^M | 1728 ^C | 2280 ^C | 1814 ^C | 0.879 | 1.190 | 0.952 | 0.200 | 1.775 |

Appendix Table A5 The statistic result of the pentosan calibration

| No. | Wavelength (nm) | | | | R | SEC* | SEP* | Bias | RPD |
|------|-------------------|-------------------|-------------------|-------------------|--------|-------|-------|-------|-------|
| 1 | 1490 ^M | | | | 0.281 | 1.310 | 0.920 | 0.252 | 1.130 |
| 2 | 1490 ^M | 2390 ^C | | | 0.736 | 0.935 | 0.782 | 0.237 | 1.330 |
| 3 | 1490 ^M | 2390 ^C | 1682 ^C | | 0.783 | 0.871 | 0.853 | 0.268 | 1.219 |
| 4 | 1490 ^M | 2390 ^C | 1682 ^C | 2340 ^C | 0.813 | 0.826 | 0.834 | 0.233 | 1.247 |
| 5 | 2122 ^M | | | | 0.202 | 1.340 | 0.985 | 0.215 | 1.056 |
| 6 | 2122 ^M | 2320 ^C | | | 0.726 | 0.950 | 0.801 | 0.193 | 1.298 |
| 7 | 2122 ^M | 2320 ^C | 1682 ^C | | 0.785 | 0.867 | 0.864 | 0.242 | 1.204 |
| 8 | 2122 ^M | 2320 ^C | 1682 ^C | 2340 ^C | 0.808 | 0.836 | 0.804 | 0.276 | 1.294 |
| 9 | 2240 ^M | | | | -0.410 | 1.240 | 0.975 | 0.303 | 1.067 |
| 10 | 2240 ^M | 2320 ^C | | | 0.747 | 0.919 | 0.827 | 0.276 | 1.258 |
| 11** | 2240 ^M | 2320 ^C | 1604 ^C | | 0.773 | 0.888 | 0.761 | 0.240 | 1.367 |
| 12 | 2240 ^M | 2320 ^C | 1604 ^C | 1652 ^C | 0.794 | 0.861 | 0.792 | 0.206 | 1.313 |
| 13 | 2284 ^M | | | | 0.527 | 1.160 | 0.742 | 0.263 | 1.402 |
| 14 | 2284 ^M | 2066 ^C | | | 0.705 | 0.979 | 0.807 | 0.247 | 1.289 |
| 15 | 2284 ^M | 2066 ^C | 2318 ^C | | 0.740 | 0.941 | 0.906 | 0.000 | 1.148 |
| 16 | 2284 ^M | 2066 ^C | 2318 ^C | 1684 ^C | 0.799 | 0.852 | 0.812 | 0.213 | 1.281 |
| 17 | 2320 ^C | | | | -0.692 | 0.984 | 0.784 | 0.190 | 1.327 |
| 18 | 2320 ^C | 1684 ^C | | | 0.793 | 0.840 | 0.833 | 0.206 | 1.248 |
| 19 | 2320 ^C | 1684 ^C | 1280 ^C | | 0.824 | 0.794 | 0.813 | 0.28 | 1.279 |
| 20 | 2320 ^C | 1684 ^C | 1280 ^C | 1734 ^C | 0.837 | 0.776 | 0.817 | 0.286 | 1.273 |
| 21 | 2456 ^M | | | | -0.405 | 1.250 | 0.904 | 0.164 | 1.150 |
| 22 | 2456 ^M | 2318 ^C | | | 0.719 | 0.960 | 0.765 | 0.153 | 1.359 |
| 23 | 2456 ^M | 2318 ^C | 1684 ^C | | 0.796 | 0.847 | 0.848 | 0.221 | 1.226 |
| 24 | 2456 ^M | 2318 ^C | 1684 ^C | 1280 ^C | 0.824 | 0.803 | 0.803 | 0.273 | 1.295 |

Appendix Table A6 The statistic result of the Lignin content calibration.

| No. | Wavelength (nm) | | | | R | SEC* | SEP* | Bias | RPD |
|------------------|-------------------|-------------------|-------------------|-------------------|--------|-------|-------|-------|-------|
| 1 | 1356 ^M | | | | 0.449 | 1.190 | 0.710 | 0.319 | 1.056 |
| 2 | 1356 ^M | 1676 ^C | | | 0.814 | 0.786 | 0.522 | 0.273 | 1.437 |
| 3 | 1358 ^M | 1676 ^C | 1240 ^C | | 0.835 | 0.755 | 0.780 | 0.467 | 0.962 |
| 4 | 1358 ^M | 1676 ^C | 1240 ^C | 2286 ^C | 0.850 | 0.732 | 0.761 | 0.484 | 0.986 |
| 5 | 1406 ^M | | | | -0.261 | 1.290 | 0.746 | 0.436 | 1.005 |
| 6 | 1406 ^M | 1674 ^C | | | 0.797 | 0.819 | 0.517 | 0.343 | 1.451 |
| 7 | 1406 ^M | 1674 ^C | 2288 ^C | | 0.842 | 0.741 | 0.53 | 0.331 | 1.415 |
| 8 | 1406 ^M | 1674 ^C | 2288 ^C | 1272 ^C | 0.854 | 0.723 | 0.51 | 0.354 | 1.471 |
| 9 | 1440 ^M | | | | 0.185 | 1.310 | 0.745 | 0.329 | 1.007 |
| 10 | 1440 ^M | 1672 ^C | | | 0.792 | 0.827 | 0.486 | 0.303 | 1.543 |
| 11 | 1440 ^M | 1672 ^C | 2430 ^C | | 0.831 | 0.764 | 0.441 | 0.298 | 1.701 |
| 12 | 1440 ^M | 1672 ^C | 2430 ^C | 2130 ^C | 0.846 | 0.742 | 0.514 | 0.243 | 1.459 |
| 13 | 1672 ^C | | | | -0.787 | 0.826 | 0.489 | 0.318 | 1.534 |
| 14 ^{**} | 1672 ^C | 2430 ^C | | | 0.833 | 0.746 | 0.442 | 0.275 | 1.697 |
| 15 | 1672 ^C | 2430 ^C | 1498 ^C | | 0.846 | 0.731 | 0.471 | 0.270 | 1.592 |
| 16 | 1672 ^C | 2430 ^C | 1498 ^C | 1238 ^C | 0.857 | 0.716 | 0.456 | 0.344 | 1.645 |
| 17 | 1920 ^M | | | | -0.249 | 1.300 | 0.737 | 0.480 | 1.018 |
| 18 | 1920 ^M | 1672 ^C | | | 0.788 | 0.833 | 0.473 | 0.305 | 1.586 |
| 19 | 1920 ^M | 1672 ^C | 1836 ^C | | 0.840 | 0.744 | 0.496 | 0.286 | 1.512 |
| 20 | 1920 ^M | 1672 ^C | 1836 ^C | 2288 ^C | 0.859 | 0.712 | 0.491 | 0.302 | 1.527 |
| 21 | 2154 ^M | | | | -0.590 | 1.080 | 0.646 | 0.948 | 1.161 |
| 22 | 2154 ^M | 1678 ^C | | | 0.798 | 0.817 | 0.491 | 0.285 | 1.527 |
| 23 | 2154 ^M | 1678 ^C | 1718 ^C | | 0.835 | 0.755 | 0.492 | 0.325 | 1.524 |
| 24 | 2154 ^M | 1678 ^C | 1718 ^C | 1892 ^C | 0.850 | 0.732 | 0.516 | 0.267 | 1.453 |

Appendix Table A6 (Cont'd)

| No. | Wavelength (nm) | | | | R | SEC* | SEP* | Bias | RPD |
|-----|-------------------|-------------------|-------------------|-------------------|--------|-------|-------|-------|-------|
| 25 | 2270 ^M | | | | 0.344 | 1.260 | 0.713 | 0.269 | 1.052 |
| 26 | 2270 ^M | 1676 ^C | | | 0.793 | 0.826 | 0.499 | 0.257 | 1.503 |
| 27 | 2270 ^M | 1676 ^C | 1718 ^C | | 0.830 | 0.765 | 0.478 | 0.324 | 1.569 |
| 28 | 2270 ^M | 1676 ^C | 1718 ^C | 2450 ^C | 0.848 | 0.737 | 0.491 | 0.314 | 1.527 |
| 29 | 2330 ^M | | | | 0.530 | 1.130 | 0.674 | 0.399 | 1.113 |
| 30 | 2330 ^M | 1676 ^C | | | 0.813 | 0.788 | 0.488 | 0.319 | 1.537 |
| 31 | 2330 ^M | 1676 ^C | 1764 ^C | | 0.841 | 0.741 | 0.501 | 0.331 | 1.497 |
| 32 | 2330 ^M | 1676 ^C | 1764 ^C | 2450 ^C | 0.854 | 0.724 | 0.512 | 0.303 | 1.465 |
| 33 | 2390 ^M | | | | -0.589 | 1.080 | 0.700 | 0.270 | 1.071 |
| 34 | 2390 ^M | 1674 ^C | | | 0.828 | 0.791 | 0.488 | 0.302 | 1.537 |
| 35 | 2390 ^M | 1674 ^C | 2430 ^C | | 0.834 | 0.758 | 0.440 | 0.279 | 1.705 |
| 36 | 2390 ^M | 1674 ^C | 2430 ^C | 1500 ^C | 0.848 | 0.737 | 0.476 | 0.261 | 1.576 |

Appendix Table A7 The statistic result of the 1%NaOH solubility value calibration.

| No. | Wavelength (nm) | | | | R | SEC* | SEP* | Bias | RPD |
|-----|-------------------|-------------------|-------------------|-------------------|--------|-------|-------|--------|-------|
| 1 | 1150 ^M | | | | 0.572 | 2.630 | 2.210 | 0.615 | 1.176 |
| 2 | 1150 ^M | 2374 ^C | | | 0.920 | 1.270 | 1.000 | 0.191 | 2.600 |
| 3 | 1150 ^M | 2374 ^C | 1200 ^C | | 0.931 | 1.200 | 0.973 | -0.024 | 2.672 |
| 4 | 1150 ^M | 2374 ^C | 1200 ^C | 1890 ^C | 0.937 | 1.170 | 0.862 | -0.026 | 3.016 |
| 5 | 1272 ^M | | | | 0.063 | 3.200 | 2.540 | 0.909 | 1.024 |
| 6 | 1272 ^M | 2374 ^C | | | 0.909 | 1.350 | 0.932 | 0.210 | 2.790 |
| 7 | 1272 ^M | 2374 ^C | 1146 ^C | | 0.927 | 1.240 | 0.939 | 0.101 | 2.769 |
| 8 | 1272 ^M | 2374 ^C | 1146 ^C | 1552 ^C | 0.933 | 1.200 | 0.857 | 0.031 | 3.034 |
| 9 | 1352 ^M | | | | 0.471 | 2.830 | 2.330 | 0.306 | 1.116 |
| 10 | 1352 ^M | 2374 ^C | | | 0.909 | 1.350 | 0.912 | 0.187 | 2.851 |
| 11 | 1352 ^M | 2374 ^C | 1148 ^C | | 0.928 | 1.230 | 0.929 | 0.025 | 2.799 |
| 12 | 1352 ^M | 2374 ^C | 1148 ^C | 1614 ^C | 0.933 | 1.200 | 0.876 | 0.078 | 2.968 |
| 13 | 1458 ^M | | | | -0.602 | 2.560 | 2.190 | 1.280 | 1.187 |
| 14 | 1458 ^M | 2374 ^C | | | 0.910 | 1.350 | 0.945 | 0.161 | 2.751 |
| 15 | 1458 ^M | 2374 ^C | 2134 ^C | | 0.929 | 1.220 | 0.809 | 0.179 | 3.214 |
| 16 | 1458 ^M | 2374 ^C | 2134 ^C | 2450 ^C | 0.937 | 1.170 | 0.775 | 0.237 | 3.355 |
| 17 | 1498 ^M | | | | 0.688 | 2.330 | 2.110 | 0.563 | 1.232 |
| 18 | 1498 ^M | 2374 ^C | | | 0.919 | 1.280 | 0.874 | 0.137 | 2.975 |
| 19 | 1498 ^M | 2374 ^C | 1656 ^C | | 0.928 | 1.230 | 0.700 | 0.256 | 3.714 |
| 20 | 1498 ^M | 2374 ^C | 1656 ^C | 2454 ^C | 0.937 | 1.160 | 0.784 | 0.228 | 3.316 |
| 21 | 1522 ^M | | | | -0.487 | 2.800 | 2.030 | 0.254 | 1.281 |
| 22 | 1522 ^M | 2374 ^C | | | 0.912 | 1.330 | 0.877 | 0.132 | 2.965 |
| 23 | 1522 ^M | 2374 ^C | 1618 ^C | | 0.931 | 1.200 | 0.851 | 0.135 | 3.055 |
| 24 | 1522 ^M | 2374 ^C | 1618 ^C | 1150 ^C | 0.936 | 1.170 | 0.874 | 0.102 | 2.975 |

Appendix Table A7 (Cont'd)

| No. | Wavelength (nm) | | | | R | SEC* | SEP* | Bias | RPD |
|------|-------------------|-------------------|-------------------|-------------------|--------|-------|-------|-------|-------|
| 25 | 1588 ^M | | | | 0.781 | 2.080 | 1.700 | 0.313 | 1.529 |
| 26 | 1588 ^M | 2374 ^C | | | 0.913 | 1.330 | 0.815 | 0.188 | 3.190 |
| 27 | 1588 ^M | 2374 ^C | 2112 ^C | | 0.928 | 1.230 | 0.781 | 0.194 | 3.329 |
| 28 | 1588 ^M | 2374 ^C | 2112 ^C | 1650 ^C | 0.936 | 1.170 | 0.883 | 0.327 | 2.945 |
| 29 | 1690 ^M | | | | 0.630 | 2.490 | 1.840 | 0.743 | 1.413 |
| 30 | 1690 ^M | 2374 ^C | | | 0.911 | 1.340 | 0.875 | 0.227 | 2.971 |
| 31 | 1690 ^M | 2374 ^C | 1146 ^C | | 0.929 | 1.220 | 0.888 | 0.107 | 2.928 |
| 32 | 1690 ^M | 2374 ^C | 1146 ^C | 2454 ^C | 0.936 | 1.180 | 0.898 | 0.067 | 2.895 |
| 33 | 1756 ^M | | | | -0.650 | 2.440 | 1.610 | 0.866 | 1.615 |
| 34 | 1756 ^M | 2374 ^C | | | 0.915 | 1.310 | 1.000 | 0.110 | 2.600 |
| 35 | 1756 ^M | 2374 ^C | 1652 ^C | | 0.931 | 0.931 | 0.830 | 0.188 | 3.133 |
| 36 | 1756 ^M | 2374 ^C | 1652 ^C | 1148 ^C | 0.935 | 1.180 | 0.910 | 0.138 | 2.857 |
| 37 | 1822 ^M | | | | 0.333 | 3.020 | 2.530 | 0.896 | 1.028 |
| 38 | 1822 ^M | 2188 ^C | | | 0.914 | 1.320 | 1.200 | 0.136 | 2.167 |
| 39 | 1822 ^M | 2188 ^C | 1546 ^C | | 0.925 | 1.250 | 1.150 | 0.112 | 2.261 |
| 40 | 1822 ^M | 2188 ^C | 1546 ^C | 2464 ^C | 0.934 | 1.190 | 1.150 | 0.064 | 2.261 |
| 41 | 1934 ^M | | | | -0.328 | 3.030 | 2.390 | 1.070 | 1.088 |
| 42 | 1934 ^M | 2186 ^C | | | 0.891 | 1.480 | 1.210 | 0.219 | 2.149 |
| 43 | 1934 ^M | 2186 ^C | 1888 ^C | | 0.930 | 1.210 | 0.987 | 0.265 | 2.634 |
| 44** | 1934 ^M | 2186 ^C | 1888 ^C | 2316 ^C | 0.945 | 1.09 | 0.702 | 0.044 | 3.704 |
| 45 | 2100 ^M | | | | 0.756 | 2.100 | 1.820 | 0.528 | 1.429 |
| 46 | 2100 ^M | 2376 ^C | | | 0.919 | 1.280 | 0.778 | 0.156 | 3.342 |
| 47 | 2100 ^M | 2376 ^C | 1994 ^C | | 0.935 | 1.170 | 0.744 | 0.148 | 3.495 |
| 48 | 2100 ^M | 2376 ^C | 1994 ^C | 1682 ^C | 0.942 | 1.120 | 0.822 | 0.177 | 3.163 |

Appendix Table A7 (Cont'd)

| No. | Wavelength (nm) | | | | R | SEC* | SEP* | Bias | RPD |
|-----|-------------------|-------------------|-------------------|-------------------|--------|-------|-------|--------|-------|
| 49 | 2188 ^M | | | | -0.889 | 1.460 | 1.170 | 0.254 | 2.222 |
| 50 | 2188 ^M | 1998 ^C | | | 0.922 | 1.260 | 1.160 | -0.035 | 2.241 |
| 50 | 2188 ^M | 1998 ^C | 2270 ^C | | 0.935 | 1.170 | 0.848 | 0.099 | 3.066 |
| 52 | 2188 ^M | 1998 ^C | 2270 ^C | 1886 ^C | 0.939 | 1.150 | 0.845 | 0.060 | 3.077 |
| 53 | 2268 ^M | | | | 0.830 | 1.790 | 1.240 | 0.001 | 2.097 |
| 54 | 2268 ^M | 2372 ^C | | | 0.913 | 1.320 | 0.768 | 0.194 | 3.385 |
| 55 | 2268 ^M | 2372 ^C | 2406 ^C | | 0.932 | 1.190 | 0.796 | 0.112 | 3.266 |
| 56 | 2268 ^M | 2372 ^C | 2406 ^C | 2192 ^C | 0.939 | 1.150 | 0.746 | 0.111 | 3.485 |
| 57 | 2334 ^M | | | | 0.781 | 2.000 | 0.760 | 1.660 | 3.421 |
| 58 | 2334 ^M | 2374 ^C | | | 0.911 | 1.340 | 0.939 | 0.877 | 2.769 |
| 59 | 2334 ^M | 2374 ^C | 1618 ^C | | 0.929 | 1.220 | 0.944 | 0.845 | 2.754 |
| 60 | 2334 ^M | 2374 ^C | 1618 ^C | 2262 ^C | 0.937 | 1.160 | 0.963 | 0.689 | 2.700 |
| 61 | 2406 ^M | | | | -0.692 | 2.310 | 1.700 | 0.298 | 1.529 |
| 62 | 2406 ^M | 2374 ^C | | | 0.909 | 1.350 | 0.970 | 0.220 | 2.680 |
| 63 | 2406 ^M | 2374 ^C | 2110 ^C | | 0.928 | 1.220 | 0.779 | 0.241 | 3.338 |
| 64 | 2406 ^M | 2374 ^C | 2110 ^C | 2370 ^C | 0.938 | 1.160 | 0.818 | 0.086 | 3.178 |
| 65 | 2446 ^M | | | | -0.370 | 2.980 | 2.550 | 0.793 | 1.020 |
| 66 | 2406 ^M | 2374 ^C | | | 0.909 | 1.350 | 0.946 | 0.205 | 2.748 |
| 67 | 2406 ^M | 2374 ^C | 2110 ^C | | 0.929 | 1.210 | 0.732 | 0.167 | 3.552 |
| 68 | 2406 ^M | 2374 ^C | 2110 ^C | 2370 ^C | 0.936 | 1.170 | 0.736 | 0.099 | 3.533 |

Appendix Table A8 The statistic result of ash content calibration.

| No. | Wavelength (nm) | | | | R | SEC* | SEP* | Bias | RPD |
|-----|-------------------|-------------------|-------------------|-------------------|--------|-------|-------|-------|-------|
| 1 | 1160 ^M | | | | 0.618 | 0.176 | 0.101 | 0.019 | 1.188 |
| 2 | 1160 ^M | 2436 ^C | | | 0.816 | 0.131 | 0.096 | 0.043 | 1.250 |
| 3 | 1160 ^M | 2436 ^C | 1962 ^C | | 0.874 | 0.112 | 0.094 | 0.051 | 1.277 |
| 4 | 1160 ^M | 2436 ^C | 1962 ^C | 1967 ^C | 0.930 | 0.085 | 0.082 | 0.035 | 1.463 |
| 5 | 1252 ^M | | | | -0.527 | 0.190 | 0.105 | 0.080 | 1.143 |
| 6 | 1252 ^M | 1962 ^C | | | 0.824 | 0.129 | 0.086 | 0.046 | 1.395 |
| 7 | 1252 ^M | 1962 ^C | 1609 ^C | | 0.903 | 0.099 | 0.089 | 0.072 | 1.348 |
| 8 | 1252 ^M | 1962 ^C | 1609 ^C | 1974 ^C | 0.930 | 0.085 | 0.076 | 0.059 | 1.579 |
| 9 | 1354 ^M | | | | 0.676 | 0.165 | 0.114 | 0.107 | 1.053 |
| 10 | 1354 ^M | 2442 ^C | | | 0.851 | 0.119 | 0.103 | 0.082 | 1.165 |
| 11 | 1354 ^M | 2442 ^C | 1962 ^C | | 0.882 | 0.108 | 0.103 | 0.081 | 1.165 |
| 12 | 1354 ^M | 2442 ^C | 1962 ^C | 1978 ^C | 0.925 | 0.089 | 0.093 | 0.080 | 1.290 |
| 13 | 1400 ^M | | | | -0.399 | 0.205 | 0.117 | 0.089 | 1.026 |
| 14 | 1400 ^M | 2010 ^C | | | 0.752 | 0.149 | 0.115 | 0.069 | 1.043 |
| 15 | 1400 ^M | 2010 ^C | 1986 ^C | | 0.856 | 0.119 | 0.102 | 0.033 | 1.176 |
| 16 | 1400 ^M | 2010 ^C | 1986 ^C | 1964 ^C | 0.930 | 0.086 | 0.084 | 0.044 | 1.429 |
| 17 | 1436 ^M | | | | 0.429 | 0.202 | 0.114 | 0.051 | 1.053 |
| 18 | 1436 ^M | 2436 ^C | | | 0.782 | 0.141 | 0.111 | 0.071 | 1.081 |
| 19 | 1436 ^M | 2436 ^C | 1346 ^C | | 0.862 | 0.116 | 0.097 | 0.059 | 1.237 |
| 20 | 1436 ^M | 2436 ^C | 1346 ^C | 1962 ^C | 0.896 | 0.104 | 0.093 | 0.060 | 1.290 |
| 21 | 1714 ^M | | | | 0.451 | 0.134 | 0.113 | 0.084 | 1.062 |
| 22 | 1714 ^M | 2440 ^C | | | 0.861 | 0.116 | 0.097 | 0.078 | 1.237 |
| 23 | 1714 ^M | 2440 ^C | 1348 ^C | | 0.892 | 0.112 | 0.091 | 0.064 | 1.319 |
| 24 | 1714 ^M | 2440 ^C | 1348 ^C | 1964 ^C | 0.874 | 0.108 | 0.087 | 0.052 | 1.379 |

Appendix Table A8 (Cont'd)

| No. | Wavelength (nm) | | | | R | SEC* | SEP* | Bias | RPD |
|------|-------------------|-------------------|-------------------|-------------------|--------|-------|-------|-------|-------|
| 25 | 1752 ^M | | | | 0.386 | 0.145 | 0.123 | 0.042 | 0.976 |
| 26 | 1752 ^M | 2440 ^C | | | 0.827 | 0.127 | 0.106 | 0.057 | 1.132 |
| 27 | 1752 ^M | 2440 ^C | 1346 ^C | | 0.890 | 0.105 | 0.093 | 0.047 | 1.290 |
| 28 | 1752 ^M | 2440 ^C | 1346 ^C | 1962 ^C | 0.917 | 0.093 | 0.093 | 0.058 | 1.290 |
| 29 | 1820 ^M | | | | 0.561 | 0.185 | 0.117 | 0.124 | 1.026 |
| 30 | 1820 ^M | 2440 ^C | | | 0.789 | 0.139 | 0.112 | 0.083 | 1.071 |
| 31 | 1820 ^M | 2440 ^C | 1348 ^C | | 0.870 | 0.113 | 0.093 | 0.058 | 1.290 |
| 32 | 1820 ^M | 2440 ^C | 1348 ^C | 1964 ^C | 0.938 | 0.081 | 0.086 | 0.032 | 1.395 |
| 33 | 1912 ^M | | | | 0.243 | 0.217 | 0.106 | 0.048 | 1.132 |
| 34 | 1912 ^M | 2436 ^C | | | 0.786 | 0.140 | 0.107 | 0.055 | 1.121 |
| 35 | 1912 ^M | 2436 ^C | 1346 ^C | | 0.863 | 0.116 | 0.093 | 0.048 | 1.290 |
| 36 | 1912 ^M | 2436 ^C | 1346 ^C | 1964 ^C | 0.892 | 0.105 | 0.095 | 0.059 | 1.263 |
| 37 | 1964 ^M | | | | -0.520 | 0.191 | 0.114 | 0.092 | 1.053 |
| 38 | 1964 ^M | 1296 ^M | | | 0.841 | 0.123 | 0.080 | 0.036 | 1.500 |
| 39 | 1964 ^M | 1296 ^M | 1974 ^M | | 0.914 | 0.094 | 0.066 | 0.026 | 1.818 |
| 40** | 1964 ^M | 1296 ^M | 1974 ^M | 1902 ^M | 0.943 | 0.077 | 0.063 | 0.042 | 1.905 |
| 41 | 2010 ^M | | | | -0.744 | 0.150 | 0.114 | 0.064 | 1.053 |
| 42 | 2010 ^M | 1986 ^C | | | 0.856 | 0.117 | 0.102 | 0.033 | 1.176 |
| 43 | 2010 ^M | 1986 ^C | 1966 ^C | | 0.925 | 0.088 | 0.080 | 0.036 | 1.500 |
| 44 | 2010 ^M | 1986 ^C | 1966 ^C | 1170 ^C | 0.939 | 0.080 | 0.067 | 0.027 | 1.791 |
| 45 | 2092 ^M | | | | 0.616 | 0.176 | 0.117 | 0.082 | 1.026 |
| 46 | 2092 ^M | 2440 ^C | | | 0.807 | 0.134 | 0.111 | 0.067 | 1.081 |
| 47 | 2092 ^M | 2440 ^C | 1348 ^C | | 0.870 | 0.113 | 0.093 | 0.059 | 1.290 |
| 48 | 2092 ^M | 2440 ^C | 1348 ^C | 1962 ^C | 0.911 | 0.096 | 0.092 | 0.059 | 1.304 |

Appendix Table A8 (Cont'd)

| No. | Wavelength (nm) | | | | R | SEC* | SEP* | Bias | RPD |
|-----|-------------------|-------------------|-------------------|-------------------|--------|-------|-------|-------|-------|
| 49 | 2120 ^M | | | | 0.660 | 0.168 | 0.114 | 0.064 | 1.053 |
| 50 | 2120 ^M | 2436 ^C | | | 0.792 | 0.138 | 0.110 | 0.066 | 1.091 |
| 51 | 2120 ^M | 2436 ^C | 1348 ^C | | 0.861 | 0.117 | 0.097 | 0.066 | 1.237 |
| 52 | 2120 ^M | 2436 ^C | 1348 ^C | 2392 ^C | 0.891 | 0.106 | 0.087 | 0.062 | 1.379 |
| 53 | 2236 ^M | | | | -0.561 | 0.185 | 0.116 | 0.058 | 1.034 |
| 54 | 2236 ^M | 2438 ^C | | | 0.808 | 0.134 | 0.109 | 0.057 | 1.101 |
| 55 | 2236 ^M | 2438 ^C | 1346 ^C | | 0.877 | 0.111 | 0.107 | 0.062 | 1.121 |
| 56 | 2236 ^M | 2438 ^C | 1346 ^C | 1758 ^C | 0.905 | 0.099 | 0.098 | 0.066 | 1.224 |
| 57 | 2296 ^M | | | | -0.534 | 0.189 | 0.108 | 0.049 | 1.111 |
| 58 | 2296 ^M | 2436 ^C | | | 0.823 | 0.129 | 0.105 | 0.054 | 1.143 |
| 59 | 2296 ^M | 2436 ^C | 1962 ^M | | 0.866 | 0.115 | 0.105 | 0.065 | 1.143 |
| 60 | 2296 ^M | 2436 ^C | 1962 ^M | 1286 ^M | 0.903 | 0.100 | 0.077 | 0.027 | 1.558 |
| 61 | 2338 ^M | | | | -0.707 | 0.158 | 0.114 | 0.080 | 1.053 |
| 62 | 2338 ^M | 1962 ^C | | | 0.793 | 0.138 | 0.110 | 0.091 | 1.091 |
| 63 | 2338 ^M | 1962 ^C | 1948 ^C | | 0.871 | 0.113 | 0.098 | 0.063 | 1.224 |
| 64 | 2338 ^M | 1962 ^C | 1948 ^C | 1344 ^C | 0.925 | 0.088 | 0.088 | 0.042 | 1.364 |
| 65 | 2436 ^M | | | | -0.780 | 0.140 | 0.110 | 0.067 | 1.091 |
| 66 | 2436 ^M | 2384 ^C | | | 0.837 | 0.124 | 0.104 | 0.052 | 1.154 |
| 67 | 2436 ^M | 2384 ^C | 1346 ^C | | 0.882 | 0.108 | 0.092 | 0.048 | 1.304 |
| 68 | 2436 ^M | 2384 ^C | 1346 ^C | 1500 ^C | 0.905 | 0.099 | 0.085 | 0.033 | 1.412 |

Appendix Table A9 The statistic result of extractive content calibration.

| No. | Wavelength (nm) | | | | R | SEC* | SEP* | Bias | RPD |
|-----|-------------------|-------------------|-------------------|-------------------|--------|-------|-------|--------|-------|
| 1 | 1180 ^M | | | | -0.252 | 0.942 | 0.838 | 0.181 | 1.089 |
| 2 | 1180 ^M | 2198 ^C | | | 0.844 | 0.529 | 0.549 | 0.190 | 1.663 |
| 3 | 1180 ^M | 2198 ^C | 2286 ^C | | 0.900 | 0.436 | 0.531 | 0.198 | 1.719 |
| 4 | 1180 ^M | 2198 ^C | 2286 ^C | 1508 ^C | 0.921 | 0.393 | 0.553 | 0.091 | 1.651 |
| 5 | 1238 ^M | | | | 0.533 | 0.823 | 0.865 | -0.012 | 1.055 |
| 6 | 1238 ^M | 2248 ^C | | | 0.840 | 0.535 | 0.612 | 0.232 | 1.492 |
| 7 | 1238 ^M | 2248 ^C | 1356 ^C | | 0.885 | 0.465 | 0.539 | 0.221 | 1.694 |
| 8 | 1238 ^M | 2248 ^C | 1356 ^C | 1502 ^C | 0.920 | 0.397 | 0.452 | 0.183 | 2.020 |
| 9 | 1332 ^M | | | | -0.281 | 0.934 | 0.911 | 0.149 | 1.002 |
| 10 | 1332 ^M | 1454 ^C | | | 0.825 | 0.557 | 0.621 | 0.009 | 1.470 |
| 11 | 1332 ^M | 1454 ^C | 1438 ^C | | 0.888 | 0.460 | 0.591 | 0.023 | 1.545 |
| 12 | 1332 ^M | 1454 ^C | 1438 ^C | 2082 ^C | 0.908 | 0.423 | 0.534 | 0.026 | 1.710 |
| 13 | 1352 ^M | | | | 0.360 | 0.908 | 0.905 | 0.164 | 1.009 |
| 14 | 1352 ^M | 2252 ^C | | | 0.847 | 0.524 | 0.515 | 0.378 | 1.773 |
| 15 | 1352 ^M | 2252 ^C | 1506 ^C | | 0.893 | 0.449 | 0.447 | 0.231 | 2.043 |
| 16 | 1352 ^M | 2252 ^C | 1506 ^C | 2286 ^C | 0.916 | 0.405 | 0.475 | 0.179 | 1.922 |
| 17 | 1450 ^M | | | | -0.745 | 0.649 | 0.688 | 0.035 | 1.327 |
| 18 | 1450 ^M | 1432 ^C | | | 0.886 | 0.457 | 0.597 | 0.018 | 1.529 |
| 19 | 1450 ^M | 1432 ^C | 1806 ^C | | 0.903 | 0.429 | 0.564 | 0.046 | 1.619 |
| 20 | 1450 ^M | 1432 ^C | 1806 ^C | 1738 ^C | 0.925 | 0.384 | 0.623 | 0.066 | 1.465 |
| 21 | 1580 ^M | | | | 0.571 | 0.799 | 0.827 | 0.103 | 1.104 |
| 22 | 1580 ^M | 2242 ^C | | | 0.832 | 0.547 | 0.652 | 0.209 | 1.400 |
| 23 | 1580 ^M | 2242 ^C | 1494 ^C | | 0.897 | 0.442 | 0.555 | 0.146 | 1.645 |
| 24 | 1580 ^M | 2242 ^C | 1494 ^C | 2192 ^C | 0.917 | 0.403 | 0.552 | 0.166 | 1.654 |

Appendix Table A9 (Cont'd)

| No. | Wavelength (nm) | | | | R | SEC* | SEP* | Bias | RPD |
|-----|-------------------|-------------------|-------------------|-------------------|--------|-------|-------|--------|-------|
| 1 | 1688 ^M | | | | 0.715 | 0.680 | 0.722 | 0.338 | 1.265 |
| 2 | 1688 ^M | 2288 ^C | | | 0.860 | 0.503 | 0.521 | 0.345 | 1.752 |
| 3 | 1688 ^M | 2288 ^C | 1506 ^C | | 0.897 | 0.440 | 0.510 | 0.184 | 1.790 |
| 4 | 1688 ^M | 2288 ^C | 1506 ^C | 1174 ^C | 0.919 | 0.399 | 0.448 | 0.184 | 2.038 |
| 5 | 1762 ^M | | | | -0.766 | 0.626 | 0.636 | 0.203 | 1.436 |
| 6 | 1762 ^M | 2132 ^C | | | 0.873 | 0.480 | 0.543 | 0.166 | 1.681 |
| 7 | 1762 ^M | 2132 ^C | 1708 ^C | | 0.905 | 0.424 | 0.514 | 0.167 | 1.776 |
| 8 | 1762 ^M | 2132 ^C | 1708 ^C | 1612 ^C | 0.925 | 0.384 | 0.436 | 0.288 | 2.094 |
| 9 | 1820 ^M | | | | 0.311 | 0.924 | 0.899 | 0.098 | 1.016 |
| 10 | 1820 ^M | 1894 ^C | | | 0.822 | 0.561 | 0.513 | 0.183 | 1.780 |
| 11 | 1820 ^M | 1894 ^C | 1986 ^C | | 0.868 | 0.496 | 0.519 | 0.014 | 1.759 |
| 12 | 1820 ^M | 1894 ^C | 1986 ^C | 1542 ^C | 0.912 | 0.414 | 0.454 | -0.012 | 2.011 |
| 13 | 1860 ^M | | | | 0.313 | 0.924 | 0.899 | 0.098 | 1.016 |
| 14 | 1860 ^M | 1466 ^C | | | 0.822 | 0.561 | 0.513 | 0.183 | 1.780 |
| 15 | 1860 ^M | 1466 ^C | 1810 ^C | | 0.868 | 0.496 | 0.519 | 0.014 | 1.759 |
| 16 | 1860 ^M | 1466 ^C | 1810 ^C | 1734 ^C | 0.912 | 0.414 | 0.454 | -0.012 | 2.011 |
| 17 | 1938 ^M | | | | -0.519 | 0.832 | 0.742 | 0.169 | 1.230 |
| 18 | 1938 ^M | 2248 ^C | | | 0.819 | 0.566 | 0.627 | 0.353 | 1.456 |
| 19 | 1938 ^M | 2248 ^C | 1458 ^C | | 0.896 | 0.443 | 0.552 | 0.171 | 1.654 |
| 20 | 1938 ^M | 2248 ^C | 1458 ^C | 1538 ^C | 0.922 | 0.393 | 0.589 | 0.128 | 1.550 |
| 21 | 2096 ^M | | | | 0.559 | 0.807 | 0.842 | 0.083 | 1.084 |
| 22 | 2096 ^M | 2246 ^C | | | 0.824 | 0.559 | 0.636 | 0.260 | 1.436 |
| 23 | 2096 ^M | 2246 ^C | 1498 ^C | | 0.902 | 0.430 | 0.515 | 0.197 | 1.773 |
| 24 | 2096 ^M | 2246 ^C | 1498 ^C | 1354 ^C | 0.920 | 0.396 | 0.444 | 0.197 | 2.056 |

Appendix Table A9 (Cont'd)

| No. | Wavelength (nm) | | | | R | SEC* | SEP* | Bias | RPD |
|-----|-------------------|-------------------|-------------------|-------------------|--------|-------|-------|--------|-------|
| 1 | 2150 ^M | | | | -0.766 | 0.625 | 0.710 | 0.101 | 1.286 |
| 2 | 2150 ^M | 1676 ^C | | | 0.850 | 0.519 | 0.529 | 0.175 | 1.726 |
| 3 | 2150 ^M | 1676 ^C | 2286 ^C | | 0.885 | 0.466 | 0.485 | 0.175 | 1.882 |
| 4 | 2150 ^M | 1676 ^C | 2286 ^C | 1504 ^C | 0.913 | 0.413 | 0.475 | 0.071 | 1.922 |
| 5 | 2250 ^M | | | | 0.817 | 0.562 | 0.608 | 0.332 | 1.502 |
| 6 | 2250 ^M | 1978 ^C | | | 0.865 | 0.496 | 0.603 | 0.287 | 1.514 |
| 7 | 2250 ^M | 1978 ^C | 1242 ^C | | 0.903 | 0.430 | 0.569 | 0.192 | 1.605 |
| 8 | 2250 ^M | 1978 ^C | 1242 ^C | 1506 ^C | 0.912 | 0.415 | 0.556 | 0.144 | 1.642 |
| 9 | 2292 ^M | | | | -0.609 | 0.776 | 0.670 | 0.176 | 1.363 |
| 10 | 2292 ^M | 1462 ^C | | | 0.868 | 0.490 | 0.519 | 0.065 | 1.759 |
| 11 | 2292 ^M | 1462 ^C | 1542 ^C | | 0.893 | 0.448 | 0.510 | 0.084 | 1.790 |
| 12 | 2292 ^M | 1462 ^C | 1542 ^C | 1148 ^C | 0.917 | 0.403 | 0.465 | 0.052 | 1.963 |
| 13 | 2328 ^M | | | | 0.615 | 0.767 | 0.798 | 0.057 | 1.144 |
| 14 | 2328 ^M | 1456 ^C | | | 0.859 | 0.505 | 0.605 | -0.020 | 1.509 |
| 15 | 2328 ^M | 1456 ^C | 2082 ^C | | 0.891 | 0.453 | 0.502 | -0.034 | 1.819 |
| 16 | 2328 ^M | 1456 ^C | 2082 ^C | 1536 ^C | 0.940 | 0.346 | 0.452 | -0.011 | 2.020 |

Appendix Table A10 The statistic result of glucose content calibration.

| No. | Wavelength (nm) | | | | R | SEC* | SEP* | Bias | RPD |
|-----|-------------------|-------------------|-------------------|-------------------|--------|-------|-------|--------|-------|
| 1 | 1200 ^M | | | | -0.610 | 1.600 | 1.350 | -0.053 | 1.074 |
| 2 | 1200 ^M | 1730 ^C | | | 0.766 | 1.310 | 1.120 | -0.146 | 1.295 |
| 3 | 1200 ^M | 1730 ^C | 2310 ^C | | 0.840 | 1.120 | 0.955 | -0.074 | 1.518 |
| 4 | 1200 ^M | 1730 ^C | 2310 ^C | 1814 ^C | 0.862 | 1.060 | 0.972 | -0.163 | 1.492 |
| 5 | 1350 ^M | | | | -0.421 | 1.830 | 1.450 | 0.268 | 1.000 |
| 6 | 1350 ^M | 2280 ^C | | | 0.762 | 1.320 | 1.010 | -0.024 | 1.436 |
| 7 | 1350 ^M | 2280 ^C | 1728 ^C | | 0.823 | 1.170 | 0.979 | -0.023 | 1.481 |
| 8 | 1350 ^M | 2280 ^C | 1728 ^C | 2312 ^C | 0.855 | 1.080 | 0.970 | -0.177 | 1.495 |
| 9 | 1494 ^M | | | | -0.655 | 1.520 | 1.170 | 0.223 | 1.239 |
| 10 | 1494 ^M | 2280 ^C | | | 0.764 | 1.320 | 1.070 | -0.070 | 1.355 |
| 11 | 1494 ^M | 2280 ^C | 1730 ^C | | 0.837 | 1.130 | 0.899 | -0.002 | 1.613 |
| 12 | 1494 ^M | 2280 ^C | 1730 ^C | 2314 ^C | 0.861 | 1.070 | 0.895 | -0.095 | 1.620 |
| 13 | 1596 ^M | | | | -0.696 | 1.450 | 1.060 | 0.245 | 1.368 |
| 14 | 1596 ^M | 2280 ^C | | | 0.800 | 1.220 | 0.996 | 0.011 | 1.456 |
| 15 | 1596 ^M | 2280 ^C | 1730 ^C | | 0.822 | 1.180 | 1.010 | -0.016 | 1.436 |
| 16 | 1596 ^M | 2280 ^C | 1730 ^C | 2310 ^C | 0.846 | 1.120 | 1.080 | -0.179 | 1.343 |
| 17 | 1784 ^M | | | | -0.566 | 1.660 | 1.130 | 0.461 | 1.283 |
| 18 | 1784 ^M | 2284 ^C | | | 0.773 | 1.290 | 1.090 | -0.036 | 1.330 |
| 19 | 1784 ^M | 2284 ^C | 1730 ^C | | 0.843 | 1.110 | 0.959 | -0.010 | 1.512 |
| 20 | 1784 ^M | 2284 ^C | 1730 ^C | 2392 ^C | 0.875 | 1.02 | 0.915 | -0.218 | 1.585 |
| 21 | 1820 ^M | | | | -0.203 | 1.970 | 1.430 | 0.392 | 1.014 |
| 22 | 1820 ^M | 2282 ^C | | | 0.774 | 1.290 | 1.100 | -0.104 | 1.318 |
| 23 | 1820 ^M | 2282 ^C | 1730 ^C | | 0.842 | 1.120 | 0.868 | -0.001 | 1.671 |
| 24 | 1820 ^M | 2282 ^C | 1730 ^C | 1760 ^C | 0.876 | 1.010 | 0.857 | -0.210 | 1.692 |

Appendix Table A10 (Cont'd)

| No. | Wavelength (nm) | | | | R | SEC* | SEP* | Bias | RPD |
|------|-------------------|-------------------|-------------------|-------------------|--------|-------|-------|--------|-------|
| 25 | 2278 ^C | | | | -0.760 | 1.310 | 1.030 | -0.028 | 1.408 |
| 26 | 2278 ^C | 1730 ^C | | | 0.820 | 1.170 | 0.901 | -0.045 | 1.609 |
| 27 | 2278 ^C | 1730 ^C | 2310 ^C | | 0.863 | 1.050 | 0.892 | -0.146 | 1.626 |
| 28** | 2278 ^C | 1730 ^C | 2310 ^C | 1814 ^C | 0.874 | 1.020 | 0.827 | -0.119 | 1.753 |
| 29 | 2338 ^M | | | | -0.669 | 1.500 | 1.020 | 0.314 | 1.422 |
| 30 | 2338 ^M | 1726 ^C | | | 0.764 | 1.320 | 0.984 | 0.264 | 1.474 |
| 31 | 2338 ^M | 1726 ^C | 2068 ^C | | 0.837 | 1.130 | 0.935 | 0.171 | 1.551 |
| 32 | 2338 ^M | 1726 ^C | 2068 ^C | 2238 ^C | 0.866 | 1.050 | 0.904 | -0.164 | 1.604 |
| 33 | 2374 ^M | | | | -0.662 | 1.510 | 1.180 | 0.522 | 1.229 |
| 34 | 2374 ^M | 2242 ^C | | | 0.835 | 1.120 | 0.866 | -0.056 | 1.674 |
| 35 | 2374 ^M | 2242 ^C | 1346 ^C | | 0.856 | 1.070 | 0.934 | -0.220 | 1.552 |
| 36 | 2374 ^M | 2242 ^C | 1346 ^C | 1262 ^C | 0.885 | 0.974 | 0.918 | -0.435 | 1.580 |

Appendix Table A11 The statistic result of xylose content calibration.

| No. | Wavelength (nm) | | | | R | SEC* | SEP* | Bias | RPD |
|------|-------------------|-------------------|-------------------|-------------------|--------|-------|-------|--------|-------|
| 1 | 1174 ^M | | | | -0.154 | 2.370 | 1.910 | 0.435 | 1.141 |
| 2 | 1174 ^M | 2052 ^C | | | 0.655 | 1.830 | 1.530 | 0.058 | 1.425 |
| 3 | 1174 ^M | 2052 ^C | 1724 ^C | | 0.717 | 1.710 | 1.480 | 0.138 | 1.473 |
| 4 | 1174 ^M | 2052 ^C | 1724 ^C | 1358 ^C | 0.756 | 1.630 | 1.690 | -0.225 | 1.290 |
| 5 | 1472 ^M | | | | -0.521 | 2.050 | 1.520 | 0.248 | 1.434 |
| 6 | 1472 ^M | 1948 ^C | | | 0.690 | 1.760 | 1.450 | 0.123 | 1.503 |
| 7 | 1472 ^M | 1948 ^C | 1720 ^C | | 0.740 | 1.650 | 1.230 | 0.255 | 1.772 |
| 8 | 1472 ^M | 1948 ^C | 1720 ^C | 2392 ^C | 0.769 | 1.590 | 1.250 | 0.113 | 1.744 |
| 9 | 1762 ^M | | | | -0.272 | 2.310 | 1.790 | 0.565 | 1.218 |
| 10 | 1762 ^M | 2048 ^C | | | 0.676 | 1.790 | 1.490 | -0.107 | 1.463 |
| 11** | 1762 ^M | 2048 ^C | 1144 ^C | | 0.736 | 1.660 | 1.210 | -0.187 | 1.802 |
| 12 | 1762 ^M | 2048 ^C | 1144 ^C | 2360 ^C | 0.766 | 1.600 | 1.500 | -0.137 | 1.453 |
| 13 | 1930 ^M | | | | -0.560 | 1.990 | 1.480 | 0.277 | 1.473 |
| 14 | 1930 ^M | 1984 ^C | | | 0.687 | 1.760 | 1.440 | 0.156 | 1.514 |
| 15 | 1930 ^M | 1984 ^C | 1754 ^C | | 0.730 | 1.680 | 1.360 | 0.192 | 1.603 |
| 16 | 1930 ^M | 1984 ^C | 1754 ^C | 2406 ^C | 0.762 | 1.610 | 1.370 | 0.148 | 1.591 |
| 17 | 2052 ^C | | | | 0.654 | 1.810 | 1.510 | 0.073 | 1.444 |
| 18 | 2052 ^C | 1724 ^C | | | 0.717 | 1.690 | 1.460 | 0.157 | 1.493 |
| 19 | 2052 ^C | 1724 ^C | 2390 ^C | | 0.751 | 1.620 | 1.520 | 0.039 | 1.434 |
| 20 | 2052 ^C | 1724 ^C | 2390 ^C | 1980 ^C | 0.777 | 1.520 | 1.410 | 0.102 | 1.546 |
| 21 | 2314 ^M | | | | -0.101 | 2.380 | 2.180 | 0.370 | 1.000 |
| 22 | 2314 ^M | 2052 ^C | | | 0.658 | 1.830 | 1.520 | 0.088 | 1.434 |
| 23 | 2314 ^M | 2052 ^C | 1726 ^C | | 0.724 | 1.700 | 1.400 | 0.097 | 1.557 |
| 24 | 2314 ^M | 2052 ^C | 1726 ^C | 1984 ^C | 0.764 | 1.610 | 1.270 | 0.082 | 1.717 |

Appendix Table A12 The statistic result of arabinose content calibration.

| No. | Wavelength (nm) | | | | R | SEC* | SEP* | Bias | RPD |
|-----|-------------------|-------------------|-------------------|-------------------|--------|-------|-------|-------|-------|
| 1 | 1238 ^C | | | | 0.778 | 0.108 | 0.102 | 0.009 | 1.373 |
| 2 | 1238 ^C | 1148 ^C | | | 0.822 | 0.099 | 0.088 | 0.019 | 1.591 |
| 3 | 1238 ^C | 1148 ^C | 2350 ^C | | 0.828 | 0.099 | 0.090 | 0.018 | 1.556 |
| 4 | 1238 ^C | 1148 ^C | 2350 ^C | 2122 ^C | 0.832 | 0.099 | 0.092 | 0.017 | 1.522 |
| 5 | 1456 ^M | | | | -0.329 | 0.162 | 0.112 | 0.035 | 1.250 |
| 6 | 1456 ^M | 1238 ^C | | | 0.786 | 0.107 | 0.097 | 0.009 | 1.443 |
| 7 | 1456 ^M | 1238 ^C | 1148 ^C | | 0.822 | 0.100 | 0.088 | 0.019 | 1.591 |
| 8 | 1456 ^M | 1238 ^C | 1148 ^C | 1400 ^C | 0.828 | 0.100 | 0.089 | 0.016 | 1.573 |
| 9 | 1622 ^M | | | | -0.579 | 0.140 | 0.107 | 0.021 | 1.308 |
| 10 | 1622 ^M | 1240 ^C | | | 0.807 | 0.103 | 0.097 | 0.010 | 1.443 |
| 11 | 1622 ^M | 1240 ^C | 1152 ^C | | 0.838 | 0.096 | 0.091 | 0.020 | 1.538 |
| 12 | 1622 ^M | 1240 ^C | 1152 ^C | 2168 ^C | 0.847 | 0.095 | 0.089 | 0.148 | 1.573 |
| 13 | 1744 ^M | | | | -0.617 | 0.135 | 0.110 | 0.020 | 1.273 |
| 14 | 1744 ^M | 1238 ^C | | | 0.787 | 0.107 | 0.099 | 0.009 | 1.414 |
| 15 | 1744 ^M | 1238 ^C | 1148 ^C | | 0.822 | 0.100 | 0.088 | 0.002 | 1.591 |
| 16 | 1744 ^M | 1238 ^C | 1148 ^C | 2350 ^C | 0.828 | 0.100 | 0.090 | 0.018 | 1.556 |
| 17 | 1805 ^M | | | | -0.055 | 0.171 | 0.130 | 0.037 | 1.077 |
| 18 | 1805 ^M | 1238 ^C | | | 0.781 | 0.108 | 0.101 | 0.009 | 1.386 |
| 19 | 1805 ^M | 1238 ^C | 2352 ^C | | 0.824 | 0.099 | 0.092 | 0.013 | 1.522 |
| 20 | 1805 ^M | 1238 ^C | 2352 ^C | 1506 ^C | 0.833 | 0.099 | 0.092 | 0.015 | 1.522 |
| 21 | 2084 ^M | | | | 0.659 | 0.129 | 0.111 | 0.016 | 1.261 |
| 22 | 2084 ^M | 1240 ^C | | | 0.795 | 0.105 | 0.104 | 0.011 | 1.346 |
| 23 | 2084 ^M | 1240 ^C | 1152 ^C | | 0.839 | 0.096 | 0.092 | 0.020 | 1.522 |
| 24 | 2084 ^M | 1240 ^C | 1152 ^C | 2168 ^C | 0.847 | 0.095 | 0.088 | 0.015 | 1.591 |

Appendix Table A12 (Cont'd)

| No. | Wavelength (nm) | | | | R | SEC* | SEP* | Bias | RPD |
|------|-------------------|-------------------|-------------------|-------------------|--------|-------|-------|-------|-------|
| 25 | 2174 ^M | | | | -0.656 | 0.130 | 0.091 | 0.013 | 1.538 |
| 26 | 2174 ^M | 1240 ^C | | | 0.823 | 0.099 | 0.093 | 0.007 | 1.505 |
| 27 | 2174 ^M | 1240 ^C | 2192 ^C | | 0.848 | 0.093 | 0.092 | 0.002 | 1.522 |
| 28 | 2174 ^M | 1240 ^C | 2192 ^C | 2370 ^C | 0.859 | 0.091 | 0.097 | 0.003 | 1.443 |
| 29 | 2266 ^M | | | | 0.553 | 0.143 | 0.093 | 0.026 | 1.505 |
| 30 | 2266 ^M | 1240 ^C | | | 0.814 | 0.101 | 0.090 | 0.012 | 1.556 |
| 31** | 2266 ^M | 1240 ^C | 1152 ^C | | 0.843 | 0.095 | 0.086 | 0.019 | 1.628 |
| 32 | 2266 ^M | 1240 ^C | 1152 ^C | 2340 ^C | 0.851 | 0.094 | 0.087 | 0.021 | 1.609 |
| 33 | 2296 ^M | | | | -0.568 | 0.141 | 0.104 | 0.025 | 1.346 |
| 34 | 2296 ^M | 1240 ^C | | | 0.804 | 0.103 | 0.094 | 0.013 | 1.489 |
| 35 | 2296 ^M | 1240 ^C | 1152 ^C | | 0.839 | 0.096 | 0.089 | 0.021 | 1.573 |
| 36 | 2296 ^M | 1240 ^C | 1152 ^C | 2170 ^C | 0.846 | 0.095 | 0.089 | 0.015 | 1.573 |
| 37 | 2330 ^M | | | | 0.644 | 0.131 | 0.113 | 0.017 | 1.239 |
| 38 | 2330 ^M | 1240 ^C | | | 0.793 | 0.106 | 0.103 | 0.011 | 1.359 |
| 39 | 2330 ^M | 1240 ^C | 1152 ^C | | 0.837 | 0.096 | 0.092 | 0.021 | 1.522 |
| 40 | 2330 ^M | 1240 ^C | 1152 ^C | 2170 ^C | 0.848 | 0.095 | 0.088 | 0.015 | 1.591 |
| 41 | 2438 ^M | | | | -0.620 | 0.135 | 0.114 | 0.020 | 1.228 |
| 42 | 2438 ^M | 1240 ^C | | | 0.783 | 0.108 | 0.104 | 0.013 | 1.346 |
| 43 | 2438 ^M | 1240 ^C | 1152 ^C | | 0.835 | 0.097 | 0.092 | 0.022 | 1.522 |
| 44 | 248 ^M | 1240 ^C | 1152 ^C | 2010 ^C | 0.851 | 0.094 | 0.098 | 0.018 | 1.429 |

Appendix Table A13 The statistic result of galactose content calibration.

| No. | Wavelength (nm) | | | | R | SEC* | SEP* | Bias | RPD |
|------|-------------------|-------------------|-------------------|-------------------|--------|-------|-------|--------|-------|
| 1 | 1186 ^M | | | | -0.332 | 0.504 | 0.397 | 0.086 | 1.008 |
| 2 | 1186 ^M | 1752 ^C | | | 0.720 | 0.375 | 0.241 | 0.040 | 1.660 |
| 3 | 1186 ^M | 1752 ^C | 1678 ^C | | 0.774 | 0.348 | 0.228 | -0.004 | 1.754 |
| 4 | 1186 ^M | 1752 ^C | 1678 ^C | 1146 ^C | 0.808 | 0.327 | 0.235 | -0.001 | 1.702 |
| 5 | 1456 ^M | | | | -0.539 | 0.450 | 0.333 | 0.158 | 1.201 |
| 6 | 1456 ^M | 1808 ^C | | | 0.712 | 0.380 | 0.244 | -0.007 | 1.639 |
| 7 | 1456 ^M | 1808 ^C | 1396 ^C | | 0.787 | 0.339 | 0.249 | -0.047 | 1.606 |
| 8 | 1456 ^M | 1808 ^C | 1396 ^C | 2130 ^C | 0.817 | 0.321 | 0.239 | -0.046 | 1.674 |
| 9 | 1612 ^M | | | | -0.465 | 0.473 | 0.314 | 0.040 | 1.274 |
| 10 | 1612 ^M | 1756 ^C | | | 0.696 | 0.389 | 0.274 | 0.028 | 1.460 |
| 11 | 1612 ^M | 1756 ^C | 2256 ^C | | 0.770 | 0.350 | 0.262 | -0.068 | 1.527 |
| 12 | 1612 ^M | 1756 ^C | 2256 ^C | 2162 ^C | 0.794 | 0.338 | 0.212 | -0.037 | 1.887 |
| 13 | 1756 ^C | | | | -0.671 | 0.396 | 0.271 | 0.003 | 1.476 |
| 14 | 1756 ^C | 2256 ^C | | | 0.770 | 0.345 | 0.263 | -0.061 | 1.521 |
| 15 | 1756 ^C | 2256 ^C | 1678 ^C | | 0.786 | 0.339 | 0.316 | -0.001 | 1.266 |
| 16 | 1756 ^C | 2256 ^C | 1678 ^C | 1146 ^C | 0.799 | 0.334 | 0.215 | -0.019 | 1.860 |
| 17 | 2080 ^M | | | | 0.353 | 0.500 | 0.338 | 0.100 | 1.183 |
| 18 | 2080 ^M | 1754 ^C | | | 0.680 | 0.397 | 0.266 | -0.003 | 1.504 |
| 19 | 2080 ^M | 1754 ^C | 2216 ^C | | 0.803 | 0.327 | 0.216 | -0.004 | 1.852 |
| 20 | 2080 ^M | 1754 ^C | 2216 ^C | 1648 ^C | 0.815 | 0.322 | 0.230 | -0.004 | 1.739 |
| 21 | 2144 ^M | | | | -0.403 | 0.489 | 0.337 | 0.107 | 1.187 |
| 22 | 2144 ^M | 1754 ^C | | | 0.732 | 0.369 | 0.286 | -0.064 | 1.399 |
| 23** | 2144 ^M | 1754 ^C | 2200 ^C | | 0.793 | 0.334 | 0.206 | 0.012 | 1.942 |
| 24 | 2144 ^M | 1754 ^C | 2200 ^C | 1820 ^C | 0.813 | 0.323 | 0.211 | -0.013 | 1.896 |

Appendix Table A13 (Cont'd)

| No. | Wavelength (nm) | | | | R | SEC* | SEP* | Bias | RPD |
|-----|-------------------|-------------------|-------------------|-------------------|--------|-------|-------|--------|-------|
| 25 | 2246 ^M | | | | 0.450 | 0.477 | 0.339 | 0.097 | 1.180 |
| 26 | 2246 ^M | 1240 ^C | | | 0.681 | 0.396 | 0.311 | 0.040 | 1.286 |
| 27 | 2246 ^M | 1240 ^C | 1564 ^C | | 0.741 | 0.368 | 0.288 | 0.030 | 1.389 |
| 28 | 2246 ^M | 1240 ^C | 1564 ^C | 2466 ^C | 0.794 | 0.338 | 0.277 | 0.015 | 1.444 |
| 29 | 2300 ^M | | | | -0.431 | 0.482 | 0.348 | 0.051 | 1.149 |
| 30 | 2300 ^M | 1242 ^C | | | 0.709 | 0.382 | 0.333 | 0.013 | 1.201 |
| 31 | 2300 ^M | 1242 ^C | 1760 ^C | | 0.747 | 0.364 | 0.301 | -0.018 | 1.329 |
| 32 | 2300 ^M | 1242 ^C | 1760 ^C | 1810 ^C | 0.773 | 0.353 | 0.298 | -0.090 | 1.342 |
| 33 | 2368 ^M | | | | 0.470 | 0.471 | 0.319 | 0.119 | 1.254 |
| 34 | 2368 ^M | 1756 ^C | | | 0.671 | 0.401 | 0.271 | 0.003 | 1.476 |
| 35 | 2368 ^M | 1756 ^C | 2200 ^C | | 0.802 | 0.328 | 0.237 | -0.005 | 1.688 |
| 36 | 2368 ^M | 1756 ^C | 2200 ^C | 1350 ^C | 0.814 | 0.323 | 0.240 | -0.008 | 1.667 |

Appendix Table A14 The statistic result of mannose content calibration

| No. | Wavelength (nm) | | | | R | SEC* | SEP* | Bias | RPD |
|-----|-------------------|-------------------|-------------------|-------------------|--------|-------|-------|-------|-------|
| 1 | 1274 ^M | | | | -0.134 | 0.247 | 0.132 | 0.064 | 1.008 |
| 2 | 1274 ^M | 1812 ^C | | | 0.454 | 0.453 | 0.127 | 0.035 | 1.047 |
| 3 | 1274 ^M | 1812 ^C | 1534 ^C | | 0.530 | 0.217 | 0.115 | 0.061 | 1.157 |
| 4 | 1274 ^M | 1812 ^C | 1534 ^C | 1764 ^C | 0.578 | 0.211 | 0.117 | 0.088 | 1.137 |
| 5 | 1364 ^M | | | | -0.184 | 0.245 | 0.133 | 0.047 | 1.000 |
| 6 | 1364 ^M | 1812 ^C | | | 0.448 | 0.226 | 0.123 | 0.046 | 1.081 |
| 7 | 1364 ^M | 1812 ^C | 1534 ^C | | 0.517 | 0.219 | 0.116 | 0.067 | 1.147 |
| 8 | 1364 ^M | 1812 ^C | 1534 ^C | 1764 ^C | 0.570 | 0.213 | 0.111 | 0.074 | 1.198 |
| 9 | 1478 ^M | | | | 0.083 | 0.248 | 0.131 | 0.068 | 1.015 |
| 10 | 1478 ^M | 1810 ^C | | | 0.454 | 0.225 | 0.120 | 0.043 | 1.108 |
| 11 | 1478 ^M | 1810 ^C | 1766 ^C | | 0.559 | 0.212 | 0.120 | 0.074 | 1.108 |
| 12 | 1478 ^M | 1810 ^C | 1766 ^C | 1392 ^C | 0.590 | 0.209 | 0.123 | 0.056 | 1.081 |
| 13 | 1722 ^M | | | | -0.354 | 0.233 | 0.129 | 0.043 | 1.031 |
| 14 | 1722 ^M | 1694 ^C | | | 0.512 | 0.217 | 0.114 | 0.055 | 1.167 |
| 15 | 1722 ^M | 1694 ^C | 2366 ^C | | 0.564 | 0.211 | 0.120 | 0.068 | 1.108 |
| 16 | 1722 ^M | 1694 ^C | 2366 ^C | 1564 ^C | 0.577 | 0.212 | 0.199 | 0.072 | 0.668 |
| 17 | 1746 ^M | | | | 0.199 | 0.244 | 0.133 | 0.055 | 1.000 |
| 18 | 1746 ^M | 2314 ^C | | | 0.486 | 0.220 | 0.110 | 0.086 | 1.209 |
| 19 | 1746 ^M | 2314 ^C | 1398 ^C | | 0.553 | 0.213 | 0.117 | 0.052 | 1.137 |
| 20 | 1746 ^M | 2314 ^C | 1398 ^C | 1446 ^C | 0.588 | 0.210 | 0.115 | 0.048 | 1.157 |
| 21 | 1820 ^C | | | | -0.440 | 0.224 | 0.124 | 0.040 | 1.073 |
| 22 | 1820 ^C | 2320 ^C | | | 0.524 | 0.215 | 0.115 | 0.062 | 1.157 |
| 23 | 1820 ^C | 2320 ^C | 1536 ^C | | 0.558 | 0.212 | 0.120 | 0.068 | 1.108 |
| 24 | 1820 ^C | 2320 ^C | 1536 ^C | 1394 ^C | 0.553 | 0.210 | 0.113 | 0.071 | 1.177 |

Appendix Table A14 (Cont'd) The statistic result of mannose content calibration

| No. | Wavelength (nm) | | | | R | SEC* | SEP* | Bias | RPD |
|------|-------------------|-------------------|-------------------|-------------------|--------|-------|-------|-------|-------|
| 25 | 2074 ^M | | | | -0.226 | 0.233 | 0.129 | 0.043 | 1.031 |
| 26 | 2074 ^M | 1694 ^C | | | 0.512 | 0.217 | 0.114 | 0.055 | 1.167 |
| 27 | 2074 ^M | 1694 ^C | 2366 ^C | | 0.564 | 0.211 | 0.120 | 0.068 | 1.108 |
| 28 | 2074 ^M | 1694 ^C | 2366 ^C | 1564 ^C | 0.577 | 0.212 | 0.119 | 0.072 | 1.118 |
| 29 | 2130 ^M | | | | -0.092 | 0.248 | 0.133 | 0.060 | 1.000 |
| 30 | 2130 ^M | 2320 ^C | | | 0.461 | 0.224 | 0.118 | 0.079 | 1.127 |
| 31 | 2130 ^M | 2320 ^C | 1810 ^C | | 0.524 | 0.218 | 0.115 | 0.062 | 1.157 |
| 32 | 2130 ^M | 2320 ^C | 1810 ^C | 1538 ^C | 0.564 | 0.214 | 0.113 | 0.077 | 1.177 |
| 33 | 2284 ^M | | | | 0.224 | 0.243 | 0.118 | 0.086 | 1.127 |
| 34 | 2284 ^M | 2428 ^C | | | 0.509 | 0.217 | 0.124 | 0.667 | 1.073 |
| 35 | 2284 ^M | 2428 ^C | 1528 ^C | | 0.544 | 0.214 | 0.119 | 0.079 | 1.118 |
| 36 | 2284 ^M | 2428 ^C | 1528 ^C | 2232 ^C | 0.585 | 0.210 | 0.123 | 0.092 | 1.081 |
| 37 | 2324 ^M | | | | -0.627 | 0.163 | 0.113 | 0.078 | 1.177 |
| 38 | 2324 ^M | 1810 ^C | | | 0.721 | 0.147 | 0.110 | 0.050 | 1.209 |
| 39 | 2324 ^M | 1810 ^C | 1536 ^C | | 0.794 | 0.131 | 0.104 | 0.070 | 1.279 |
| 40** | 2324 ^M | 1810 ^C | 1536 ^C | 2358 ^C | 0.805 | 0.129 | 0.104 | 0.068 | 1.279 |
| 41 | 2358 ^M | | | | -0.484 | 0.184 | 0.130 | 0.041 | 1.023 |
| 42 | 2358 ^M | 2282 ^C | | | 0.672 | 0.157 | 0.113 | 0.076 | 1.177 |
| 43 | 2358 ^M | 2282 ^C | 1810 ^C | | 0.749 | 0.143 | 0.105 | 0.033 | 1.267 |
| 44 | 2358 ^M | 2282 ^C | 1810 ^C | 1340 ^C | 0.805 | 0.130 | 0.117 | 0.085 | 1.137 |

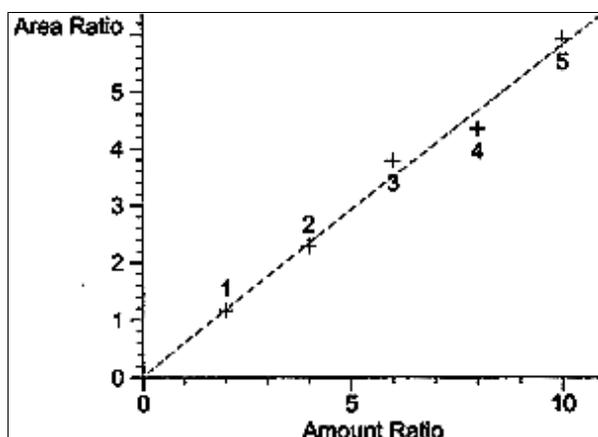
M = Selected by manual, C = Selected by computer, R = Multiple correlation coefficient of calibration, SEC = Standard error of calibration, SEP = standard error of prediction,

Bias = The average of difference between actual value and NIR value

* Unit (%W/W), ** selected calibration

APPENDIX B

Appendix Figure B1 The gas chromatography calibration curve of standard monosaccharide.



Glucose calibration

Correlation = 0.99597

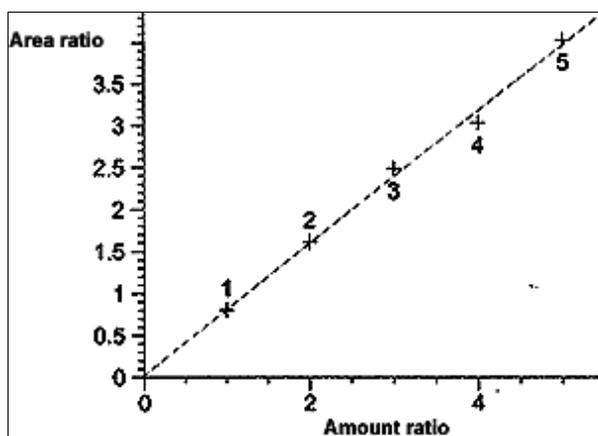
Formular : $y = mx+b$

: $m = 5.80298e-1$

: $b = 1.33285e-2$

: $x = \text{Amount ratio}$

: $y = \text{Area ratio}$



Xylose contents

Correlation = 0.99844

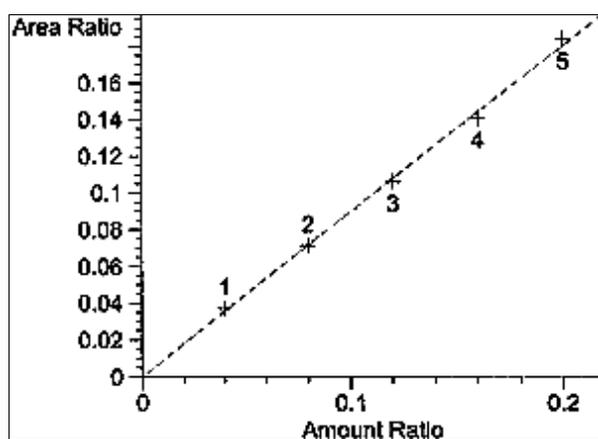
Formular : $y = mx+b$

: $m = 7.91460e-1$

: $b = 1.46491e-2$

: $x = \text{Amount ratio}$

: $y = \text{Area ratio}$



Arabinose contents

Correlation = 0.99936

Formular : $y = mx+b$

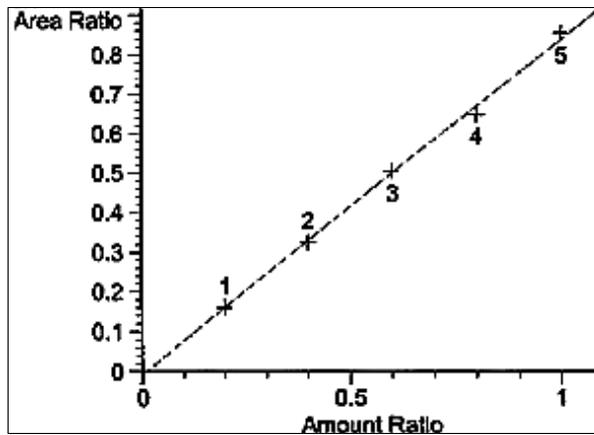
: $m = 9.02742e-1$

: $b = -5.36350e-4$

: $x = \text{Amount ratio}$

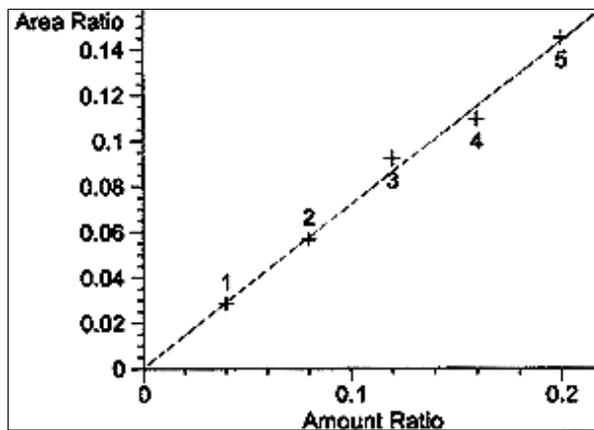
: $y = \text{Area ratio}$

Appendix Figure B1 (Cont'd)



Galactose contents

Correlation = 0.99916

Formular : $y = mx+b$: $m = 8.43912e-1$: $b = -7.80671e-3$: $x = \text{Amount ratio}$: $y = \text{Area ratio}$ 

Mannose

Correlation = 0.99762

Formular : $y = mx+b$: $m = 7.18594e-1$: $b = 2.80305e-4$: $x = \text{Amount ratio}$: $y = \text{Area ratio}$

CURRICULUM VITAE

NAME : Miss Nattaporn Suttiwijitpukdee
BIRTH DATE : May, 15th 1980
BIRTH PLACE : Suphanburi, Thailand
EDUCATION : B.Sc.(Chemistry) Thammasat Univ. (2002)
POSITION/TITLE : Temporary Staff
WORK PLACE : Kasetsart Agricultural and Agro-industrial
Product Improvement Institute (KAPI)
POSTER PRESENTATION AT CONFERENCE : Determination of chemical component in
Eucalyptus camaldulensis by Near Infrared
Spectroscopy
: Nondestructive Quality Determination of
Eucalyptus Wood Powder using Near Infrared
Spectroscopy
SCHOLARSHIP/AWARDS : NSTDA Scholarship from Thammasat Univ.
1998-2001