THESIS APPROVAL

GRADUATE SCHOOL, KASETSART UNIVERSITY

Master of Science (Forestry)

DEGREE

Forest Products	Forest Products
FIELD	**DEPARTMENT**

TITLE: Assessment of Wood Density, Syringyl/Guaiacyl Ratio of Lignin and Fiber Characteristics from Eucalyptus Hybrids Using Near Infrared Spectroscopy

NAME: Mr. Worapong Athihirunwong

THIS THESIS HAS BEEN ACCEPTED BY

THESIS ADVISOR

(Assistant Professor Pratuang Puthson, Dr.rer.nat.)

COMMITTEE MEMBER

(Assistant Professor Marisa Arunchaiya, Ph.D.)

DEPARTMENT HEAD

(Assistant Professor Sawitree Psuttipiched, D.Tech.Sc.)

APPROVED BY THE GRADUATE SCHOOL ON

DEAN

(Associate Professor Gunjana Theeragool, D.Agr.)

THESIS

ASSESSMENT OF WOOD DENSITY, SYRINGYL/GUAIACYL RATIO OF LIGNIN AND FIBER CHARACTERISTICS FROM EUCALYPTUS HYBRIDS USING NEAR INFRARED SPECTROSCOPY

WORAPONG ATHIHIRUNWONG

A Thesis Submitted in Partial Fulfillment of
The Requirements for the Degree of
Master of Science (Forestry)
Graduate School, Kasetsart University
2011

Worapong Athihirunwong 2011: Assessment of Wood Density, Syringyl/Guaiacyl Ratio of Lignin and Fiber Characteristics from Eucalyptus Hybrids Using Near Infrared Spectroscopy. Master of Science (Forestry), Major Field: Forest Products, Department of Forest Products. Thesis Advisor: Assistant Professor Pratuang Puthson, Dr.rer.nat. 118 pages.

Wood density, syringyl/guaiacyl (S/G) ratio of lignin and fiber characteristics belong to the indicators of wood quality that are much related to pulping and papermaking properties. The feasibility of near infrared (NIR) spectroscopy for the non-destructive evaluation of these properties from five eucalyptus hybrids planted in Thailand was investigated. Acquisition of NIR spectra (1000-2500 nm) was obtained from 40-60 mesh wood powder using a standard closed-cup for measuring. The best NIR models of wood basic density, S/G ratio and fiber length were obtained from partial least squares (PLS) regression analysis, whereas that of fiber coarseness obtained from principal component regression (PCR) analysis. The ratio of performance to deviation (RPD) of wood basic density (RPD = 3.13) and S/G ratio of lignin (RPD = 2.67) were fulfill the requirements of AACC Method 39-00 for screening in tree breeding programs (RPD ≥ 2.5). The RPDs of NIR models for fiber length and fiber coarseness (RPD = 1.83 and 1.97 respectively) were not good enough to use for screening in tree breeding programs. Further investigation in the relationships of these fiber characteristic properties and NIR spectra is required.

_______________________ _______________________ ___ / ___ / ___
Student's signature Thesis Advisor's signature

ACKNOWLEDGEMENTS

I gratefully thank and deeply indebted to Asst. Prof. Dr. Pratuang Puthson my thesis advisor for advice, encouragement and valuable suggestion for completely writing of thesis. I sincerely thank Asst. Prof. Dr. Marisa Arunchaiya my thesis committee; and also Prof. Dr. Satoru Tsuchikawa, Prof. Dr. Manfred Schwanninger, and Assoc. Prof. Dr. Anupun Terdwongworakul for their valuable comments and suggestions.

I gratefully thank Siam Forestry Company (SCG Paper), Kanchanaburi, Thailand and Mr. Nuttawut Muangthong for kindly providing eucalyptus hybrids for using in this study. I gratefully thank Mrs. Warunee Thanapase, Dr. Sumaporn Kasemsumran, and Ms. Jiraporn Anusornwongchai from SRU-NIR research unit, Kasetsart Agricultural and Agro-Industrial Product Improvement Institute (KAPI) for kindly support equipment of Fourier transform near infrared spectrophotometer; and also Mahaphan Fiber Cement Company Ltd. (Public), Lopburi, Thailand for kindly support equipment of Kajaani FS300 fiber analyzer. I am greatly indebted to Prof. Dr. Jürgen Odermatt, Mr. Andreas Klingberg, and Ms. Cornelia Hamann from Center of Wood Science and Technology, University of Hamburg, Hamburg, Germany for kindly providing the analysis of syringyl/guaiacyl ratio of wood lignin.

This research was supported by SRU-NIR research unit, Kasetsart Agricultural and Agro-Industrial Product Improvement Institute (KAPI), Kasetsart University, Thailand and was partially funded by the Capacity Building for KU Students on Internationalization project (Fiscal Year B.E. 2553), International Affair Division together with Faculty of Forestry, Kasetsart University, Thailand.

Finally, I especially appreciate my parents, my aunts, my sisters and brothers for their continuing encouragements during my graduate study.

Worapong Athihirunwong
May 2011

TABLE OF CONTENTS

LIST OF TABLES

Table **Page**

Appendix Table

LIST OF TABLES (Continued)

LIST OF TABLES (Continued)

Appendix Table **Page**

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF ABBREVIATIONS

NIR = Near infrared
CAL = Calibration set
VAL = Validation set
PCA = Principal component analysis
PCR = Principal component regression
PLSR = Partial least square regression
SG smoothing = Savitzky-Golay smoothing
MSC = Multiplicative scatter correction
MN = Maximum normalization
VN = Vector normalization
LBC = Linear baseline correction
SNV = Standard normal variate
1st Der = First derivative
2nd Der = Second derivative
r = Correlation coefficient
R^2 = Coefficient of determination
SD = Standard deviation
SEC = Standard error of calibration
SEP = Standard error of prediction (or validation)
SECV = Standard error of cross-validation
RMSEC = Root mean square error of calibration
RMSEP = Root mean square error of prediction (or validation)
RMSECV = Root mean square error of cross-validation
RPD = Ratio of standard error of prediction (or calibration) to sample standard deviation

ASSESSMENT OF WOOD DENSITY, SYRINGYL/GUAIACYL RATIO OF LIGNIN AND FIBER CHARACTERISTICS FROM EUCALYPTUS HYBRIDS USING NEAR INFRARED SPECTROSCOPY

INTRODUCTION

Eucalyptus species is utilized as the main fibrous raw material for the pulp production in Thailand due to its high growth rate. In order to improve pulpwood quality, many research groups established tree breeding programs, and one of them is focused on eucalyptus hybrids. Basically, the parameters to indicate the pulpwood quality apart from growth rate and insect resistance are wood properties which are wood density, fiber characteristics, and wood chemical compositions. Wood density is related to the pulp yield. Fiber characteristics, such as length and coarseness, are also important selection criteria because it expressed the papermaking properties of pulpwood fibers (Horn and Setterholm, 1990). In addition, the syringyl/guaiacyl (S/G) ratio of lignin is affected on pulping performance (Tsutsumi, 1995). High S/G ratio, lignin is more easily delignified which results higher pulp yield and less alkali consumption (Collin *et al.*, 1990; Guerra *et al.*, 2008; Del Río *et al.*, 2005).

To assess their pulpwood parameters, many standard conventional methods are required. These mostly destructive methods are costly, time consuming, require hazardous chemicals, tiresome and they are not sufficiently applicable for routine or online pulp quality control in factory. Alternatively, near infrared (NIR) spectroscopy coupled with chemometrics, which is a simple, powerful, and sensitive non-destructive tool, can provide the information of chemical and physical properties of lignocellulosic materials (Siesler *et al.*, 2002; Ozaki *et al.*, 2007). NIR absorption bands are derived from the overtones and combinations of C-H, N-H, O-H and C=O vibrations (Workman and Weyer, 2008). By this spectroscopic means, not only the chemical structure modification but also the physical and morphological changes of organic substances can be characterized.

The chemometrics is a truly interdisciplinary science that combines technical mathematics, statistics and information science in order to extract relevant information from NIR spectra (instrumentally measured) with the wood properties determined by the conventional method (laboratory analysis) is made through regression techniques (modeling) and makes it possible to predict these traits on unknown samples. Usually, multivariate statistical method such as principal component analysis (PCA), principal component regression (PCR), and partial least square regression (PLSR) are used for the evaluation of NIR spectra and for the calculation of prediction models.

As the research question that the traditional methods are unsuitable use when large screening of tree improvement programs is necessary. No any methods can provide all analysis of wood properties at the same time. Therefore, a suitable method to assess a number of the key wood properties in a large number of samples is needed to research and develop. The purpose of this study is to investigate the feasibility of NIR spectroscopy technique to assess wood density, S/G ratio of lignin and fiber characteristics of eucalyptus hybrids in Thailand.

OBJECTIVES

1. To study wood basic density, syringyl/guaiacyl ratio of lignin, fiber length and fiber coarseness from eucalyptus hybrids

2. To develop the NIR spectroscopic method for assessment of wood basic density, syringyl/guaiacyl ratio of lignin, fiber length and fiber coarseness from eucalyptus hybrids

LITERATURE REVIEW

1. Eucalyptus Species

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. The genus eucalyptus with more than 500 species is regarded by some taxonomists as several subgenera or perhaps even genera. The ranking of the ten most important eucalyptus species around the world include: *E. grandis, E. camaldulensis, E. tereticornis, E. globulus, E. urophylla, E. viminalis, E. saligna, E. deglupta, E. exserta* and then either *E. citriodora, E. paniculata or E. robusta*. Of these the first four are by far the most important. All these species are in the subgenus *Symphyomyrtus*, except *E. citriodora* which is in the subgenus *Corymbia* (Davidson, 1996).

The major eucalyptus species used in Thailand is *E. camaldulensis*, which mostly produces woodchips for pulp and paper industry. Tree improvement programs of eucalyptus species in Thailand are developed by not only government organization, but it also conducted by private companies. The main objective of eucalyptus breeding programs for pulp and paper production is to produce varieties of trees with high levels of cellulose with the least possible amount of lignin. However, trees with these characteristics can be a major problem in the field, because of their fragility that makes them susceptible to breaking. Moreover, a monoculture plantation of a single eucalyptus species usually causes an increase in plant disease. Therefore, tree breeding program is a tool to eliminate these weak points and provides a number of sustainable pulpwood with high quality.

2. Wood Lignin

Lignin exists as one of the essential wood components, ranging in amount from 10% to 30%. It is through that lignin is a polymer formed by the enzymatic dehydrogenation of phenylpropanes followed by radical coupling (Sakakibara and Sano, 2001). Macromolecule of lignin plays a vital role in providing mechanical support to bind plant fibers together. Lignin also decreases the permeation of water through the cell walls of the xylem, thereby playing an intricate role in the transport of water and nutrients. Although lignin is necessary to tree, it is undesirable in most chemical papermaking fibers and is removed by pulping and bleaching processes (Ragauskas, 2007).

2.1 Monomers of lignins

Softwood lignin is composed mainly of guaiacyl units originating from the predominant precursor, *trans*-coniferyl alcohol, while hardwood lignin is composed of both guaiacyl and syringyl units derived from *trans*-coniferyl and *trans*-sinapyl alcohols, respectively. Grass lignin contains *p*-hydroxyphenyl units derived from *trans*-*p*-coumaryl alcohol, besides units originating from the foregoing two precursors. However, strictly speaking, almost all plants consist more or less of all three units, namely, guaiacyl, syringyl, and *p*-hydroxyphenyl moieties (Figure 1) (Sakakibara and Sano, 2001).

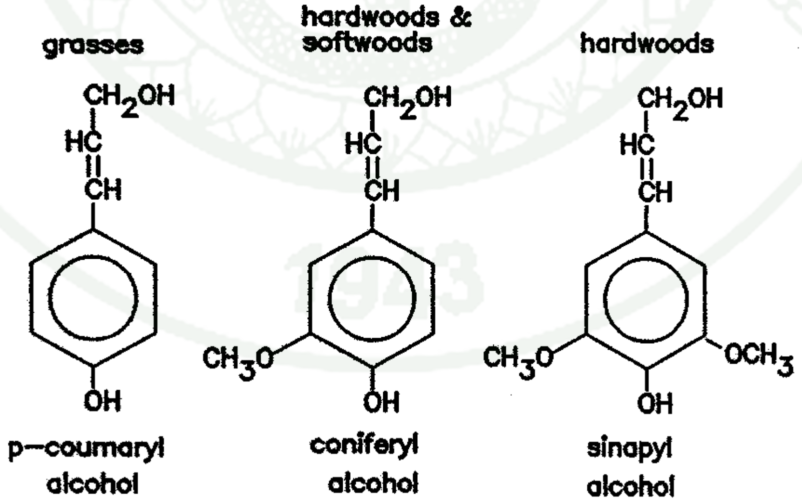


Figure 1 Lignin precursors for plants

Source: Biermann (1996)

2.2 Linkages of lignin polymer

The present of many complex carbon-to-carbon linkages between the units makes it difficult to degrade the polymer to low-molecular weight fragments. Furthermore, it has not yet been possible to isolate all parts of the lignin completely from plant tissue without engendering structural changes. These characteristics make it hard to elucidate the chemical structure of lignin. However, the dominant structures in lignin have been elucidated as the methods for identification of the degradation products and for the synthesis of model compounds have improved. The results from these numerous studies have yielded what is believed to be an accurate representation of the structure of lignin. Examples of elucidated structural features of lignin include the dominant linkages between the phenylpropane units and there abundance, as well as the abundance and frequency of some functional groups. Linkages between the phenylpropane units and the various functional groups on these units give lignin a unique and very complex structure (Sakakibara and Sano, 2001; Ragauskas, 2007).

Sakakibara and Sano (2001) pointed out that the information on the degradation products from lignin could provide direct evidence about lignin structure. For this purpose, the following procedures for bringing about lignin degradation are effective: catalytic hydrogenolysis, controlled hydrolysis, and degradation by thioacetolysis or thioacidolysis followed by reduction. As complementary procedures, oxidations of lignin with $KMnO_4$, nitrobenzene under alkaline conditions, and acidolysis are available. Summarized in Figure 2 are the main types of lignin structural units which were obtained from the various lignin degradation products. The most important linkage types in the lignin molecule are β-O-4 (**B**), and then β -5 (**D**), 5-5 (**E**), β -1 (**C**), and α-O-4 (**A**).

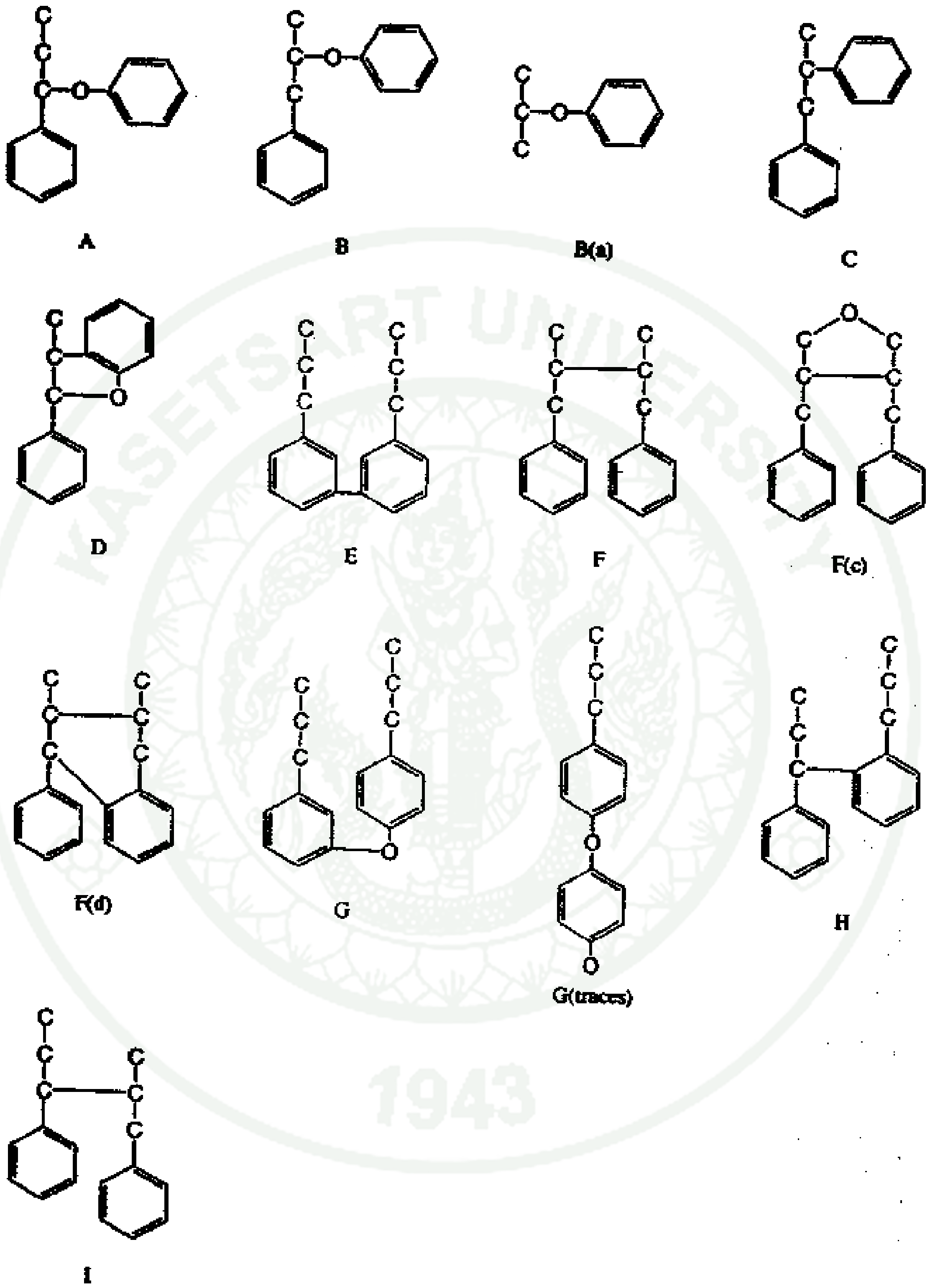


Figure 2 Typical linkage units in lignin

Source: Sakakibara and Sano (2001)

Furthermore, the quantitative data for various functional groups and linkage types in lignin are also essential to understanding lignin structure. Ultraviolet (UV), infrared (IR), and nuclear magnetic resonance (NMR) spectroscopic techniques, particularly when used in conjunction with chemical modification, have contributed to estimating the frequencies of functional groups and linkage types of lignin. Frequencies of the functional groups and the typical linkage units in birch milled wood lignin and beech lignin are collected in Table 1 (Sakakibara and Sano, 2001).

Table 1 Structural units (per 100 C_6C_3 units) of birch MWL and beech lignin

Units	Birch			Beech
	G	S	Total	Total
A: α-O-4 (open)	-	-	6	-
B: β-O-4	22 – 28	34 – 39	60	65[b]
β-O-4[a]	-	-	2	-
C: β-1	-	-	7	15
D: β-5	-	-	6	6
E: 5-5	4.5	-	4.5	2.3
F: β-β	-	-	3	5
β-β[c] d	-	-	-	2
β-β and β-6[d]	-	-	-	0.5
G: 4-O-5, 4-O-1	1	5.5	6.5	1.5
H: C(α)-2, C(α)-2	1 – 1.5	0.5 – 1	1.5 – 2.5	-
I: α–β	-	-	-	2.5

Remark: [a] In glyceraldehydes-b-aryl ether.

[b] A + B

[c] In dibenzyltetrahydrofuran units.

[d] In tetralin units.

Source: Sakakibara and Sano (2001)

2.3 Model of lignin structure

For the structural model of hardwood lignin, Nimz proposed a constitutional scheme for beech lignin on the basis of the result from mild hydrolysis and thioacetic acid degradation of beech lignin (Figure 3). This structure model consists of 25 phenylpropane units containing 14 guaiacyl, 10 syringyl, and one *p*-hydroxyphenyl moiety (Figure 4), of which six units can some extent be replaced by the diglinol units enclosed in the brackets. The models give a representative section from a beech lignin molecule 10-20 times larger, in which the 10 different bond types are randomly distributed. The formula of this structural model is calculated as $C_9H_{7.16}(OCH_3)_{1.36}$ (Sakakibara and Sano, 2001).

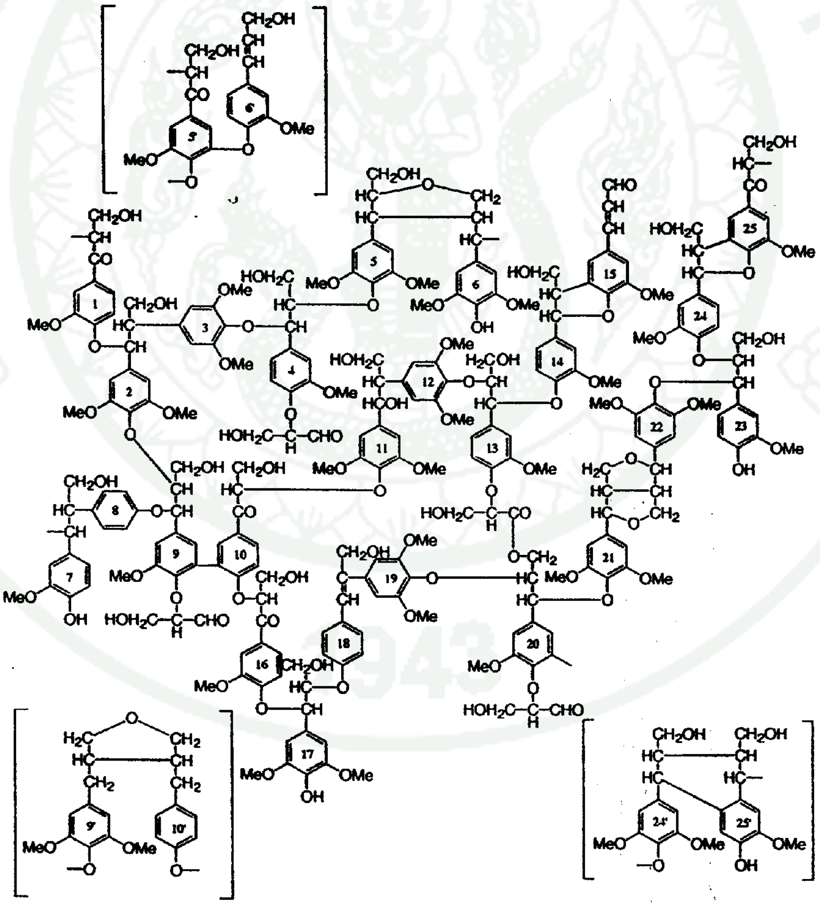


Figure 3 A structure model for beechwood lignin

Source: Sakakibara and Sano (2001)

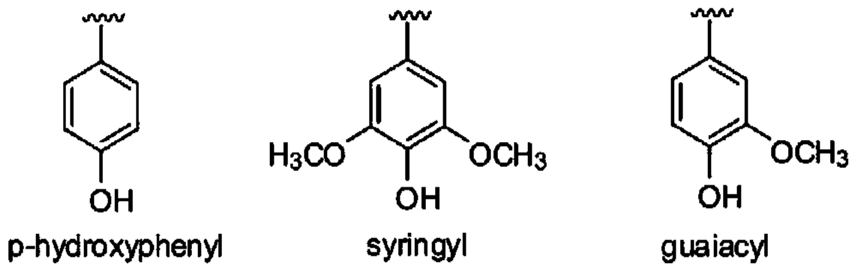


Figure 4 Constituent aromatic unit of lignin

Source: Nunes *et al.* (2010)

The question of whether the association between lignin and carbohydrates is a physical or chemical in nature has been debated for a long time and has not yet been resolved. Numerous studies have also suggested that covalent linkages between lignin and hemicelluloses exist in native wood. These structures are typically referred to as lignin-carbohydrate complexes or simply LCCs. The lignin is covalently bound to the hemicellulose which, in turn, is bound to cellulose through extensive hydrogen bonding. Lignin-carbohydrate complexes could be very important when considering pulping of wood since lignin is chemically bound to the cellulose (Ragauskas, 2007).

2.4 Importance of lignin on pulp production

Lignin is an undesirable component in the conversion of wood into pulp and paper. Lignin removal is a major step in the papermaking process. Lignin content is an important determinant with respect to cellulose fiber extraction from wood. Lignin subunit composition influences cellulose accessibility. Tree breeders are thus seeking ways to reduce extractive content and/or lignin content or modify the monomeric composition to improve pulp manufacturing. The presence of methoxylated syringyl (S) units facilitates chemical delignification during pulp manufacturing but this is not the only structural parameter which affects kraft cooking (Baillères *et al.*, 2002).

At the present time, kraft pulping and bleaching technology requires a kappa number of about 15-20 for the unbleached pulp. Bleaching is used to remove the remaining 3-5 % lignin from the fibers and transform the pulp color from brown to white (Bose *et al*., 2009).

Lignin rich in G units has relatively more carbon-carbon bonds than lignin rich in S units because the aromatic C-5 position of G units is free to make linkages. Because of β-aryl ether linkage dominance (48 percent in spruce lignin and 60 percent in birch lignin) and great reactivity during pulping and bleaching processes, the ether β-O-4 linkages are the target of the kraft delignification process, whereas the carbon-carbon bonds (β−β, β−1, β-5 and 5-5) are more resistant to chemical degradation. As a consequence, wood lignins essentially made of G units are less susceptible to kraft delignification (Chen *et al*., 2001; Kuroda *et al*., 2002).

Almost all tree breeding operation in the global pulp and paper industries are trying to selected hardwoods with high S/G ratios. Approximately two-thirds of all virgin wood pulp (recycled fibers excluded) is produced by the kraft process that uses NaOH and Na_2S to extract at least 90 percent of the lignin. The rate of this process is believed to be enhanced by a high S/G ratio. The evidence is convincing enough that reactive lignin S units are extracted from the cell wall at higher rate than lignin G units during the rapid or bulk phase of kraft pulping. However, there is no convincing evidence that the initial S/G ratio of the native lignin has any significant effect on delignification rate in the residual phase of the kraft pulping (Bose *et al.*, 2009).

Brun *et al*. (n.d.) summarized that the application of biotechnology in forestry can prove highly beneficial from both the industrial and environmental perspectives. For the case of eucalyptus plantation, using biotechnology to reduce total lignin or alter the composition of lignin in eucalyptus trees, the final yield in pulp mills can be greatly enhanced. These traits will also lead to reductions in chemicals and energy needed for the process. These direct benefits can also translate into lessened environmental impacts from the production process and/or reduced plantation areas. Many experimental results have already demonstrated lignin reductions and increased S/G ratio in research-scale tests. ArborGen (a global leader in the research, development, and commercialization of applications in biotechnology, Campinas, SP-Brazil) has obtained results from early wood quality analysis, showing that the S/G ratio can be significantly increased by the introduction of a gene that encodes a key enzyme in syringyl lignin biosynthesis.

The partial research of Collins *et al*. (1990) investigated that the lignin content and wood density of ten different hardwoods correlated to the alkali requirement for kraft delignification to kappa number approximately 20. They concluded that irrespective of lignin content and wood density of the four species with the highest S/G ratios consumed less alkali than the six species with lower S/G ratios. Species of Eucalyptus that show high S/G ratio, such as *E. globulus*, tend to also contain lower overall lignin contents and higher basic densities. Furthermore, they also studied with *E. tereticornis* and *E. deglupta* indicated that wood having lignin with high S/G ratio requires less effective alkali to reach a given kappa number and produce higher yields (Bose *et al*., 2009).

There are a number of research results that demonstrate the efficiency of transgenic trees in delivering higher potential industrial yield. In one of particular case; kraft pulping of tree trunks from poplar lines improved through biotechnology, Pilate *et al*. (2002) showed that the lines with reduced lignin had improved processing characteristics, thereby allowing for easier delignification, using smaller amounts of chemicals, while yielding more high-quality pulp.

An expected penalty of increasing S/G ratio is the slightly decrease in black liquor heating value because syringyl lignin contains less carbon than guaiacyl lignin. However, the higher pulping yield obtained with high S/G ratio may compensate for such loss since less carbohydrate loss mean black liquor with higher heating value. Furthermore, the lower effective alkali demand required for cooking high S/G ratio lignin wood decrease causticizing loads and decrease the white liquor production costs. Finally, wood containing a higher S/G ratio leads to kraft pulps containing higher S/G ratio as well. Pulps having higher S/G ratio consume less chlorine dioxide to reach full brightness (Brun *et al.*, n.d.).

3. Wood Basic Density

Density is the ratio of the mass of a quantity of a substance to its volume and is expressed in terms of weight per unit volume. TAPPI standard T 258-om02: basic density and moisture content of pulpwood defined the determination of basic density is that it obtained by dividing the oven-dry weight by the green (maximum swollen) volume (Technical Association of the Pulp and Paper Industry, 2002b). And many technologies consider the density of wood is the best expressed only on the basis of oven-dry weight and maximum volume, because at moisture contents above the fiber-saturation point the maximum volume is essentially constant.

The basic density of wood is one of the most important selection criteria in tree breeding program. It is a critical timber quality trait for the production of pulp, paper and sawn timber, and also used as excellent indicator of the mechanical and other physical properties of wood. It is the most useful index of the suitability of wood for many end-product uses (Muneri and Raymond, 2001; Seca and Domingues, 2006; Tsuchikawa, 2007).

Furthermore, the wood density is also important factor affecting pulp quality. An increase in density will increase pulp yield and tear index, but will decrease tensile, burst, apparent density and stretch. The increase in pulp yield with increased density is caused by the higher concentration of cellulose in latewood and the additional porosity of the earlywood zone. Differences in magnitude of porosity between latewood and earlywood result in uneven rates of liquor penetration and thus pulp yield, especially for shorter cooking periods. Latewood can exhibit 2 to 7 percentage points higher yield than earlywood, setting the boundaries for pulp yield variation. As a result, density variation becomes a primary factor for most pulp and paper properties (Via *et al.*, 2005).

4. Fiber Characteristics

According to TAPPI standard T 234-cm02: coarseness of pulpwood fibers defined the coarseness as the weight per unit length of fiber expressed as milligrams per 100 meter (or decigrams per 10 kilometer) and called a decigrex, abbreviated to “dg” (Technical Association of the Pulp and Paper Industry, 2002a). The coarseness of fibers has an important effect on many properties of pulpwood fiber. For example, if other factors are equal, a finer pulp will give a stronger, smoother and better folding paper. On the other hand, if assuming constant fiber diameter, coarser fibers generally have thicker cell walls, are stiffer and more inflexible, resist collapse during papermaking and, therefore, form bulkier, more porous and rougher sheets (Muneri and Raymond, 2001).

The physical properties of sheets made from pulpwood fiber are very much dependent upon fiber morphological characteristics. Fiber length can be an important factor in such hardwood sheet properties as bursting and tensile strength, tearing strength, and stretch. Sheet properties dependent upon fiber-to-fiber bond formation for strength development are dependent upon a combination of fiber length and cell wall thickness. These morphological characteristics combined, reported as a ratio of length to thickness (L/T), can be used to assess the potential of fibers for developing strength properties such as bursting and tensile strength. Without question, sheet properties are dependent upon the morphological characteristics of the pulp fiber. Processing variables, such as beating, will affect sheet strength. However, everything else being equal, sheet properties are most influenced by the morphological characteristics of the fibers (Horn and Setterholm, 1990).

5. Pyrolysis GC/MS

5.1 Introduction to pyrolysis GC/MS

The characterization of relatively complex polymers is usually carried out by means of coupled techniques because sometimes a single technique is not enough to elucidate their structures. Pyrolysis of polymers is an old technique used many years ago to identify materials by their vaporized decomposition products. The coupling of this sample method with a powerful identification technique, such as infrared (IR) spectroscopy or, often, mass spectrometry (MS), has demonstrated its utility for the analysis of polymeric materials and, mainly, for the characterization of their degradation products (Migallon and Marín, 2001; Wolf *et al*., 1980).

The pyrolysis of a sample can be defined as the conversion of a substance into other(s) with lower-molecular-weights by the action of heat. In the case of organic polymers, the most interesting data we can obtain from the fragments of pyrolysis are the composition of the volatiles produced and of the nonvolatile residues, changes in molecular weight as a function of the temperature, and evolution energies of the degradation reactions (Migallon and Marín, 2001; Wolf *et al*., 1980).

In order to obtain these data, the most often applied approach has been coupling of gas chromatography (GC) with MS. The pyrolysis products are first separated in the column and then the separated fragments are directly analysed in the mass spectrometer. Therefore, it is possible to obtain reliable and reproducible result with a single run and in a relatively short time of analysis (Migallon and Marín, 2001; Wolf *et al*., 1980).

Pyrolysis GC/MS, however, presents some limitations associated with many factors, such as the complexity of the chemical reactions involved in pyrolysis, the incomplete separation of the degradation products, and, sometimes, poor peak identification and interpretation of chromatograms. However, some recent developments have permitted the application of this technique to a broader range of samples. Those developments are principally the use of highly specific pyrolysis device, highly efficient capillary columns, and MS libraries with a wider range of spectra. This fact has permitted the use of the technique not only for polymer identification but also for structural characterization of polymers and blends (Migallon and Marín, 2001; Wolf *et al.*, 1980).

An important aspect of analytical pyrolysis is the production and detection of the thermal fragments containing essential structural information, in order to provide insight, at least partially, into the original structures of the compounds present in the chemical matrix. Larger fragmentation products are the more informative among all fragments. Therefore, there is a current tendency in analytical pyrolysis-GC/MS to preserve and detect higher-molecular-weight fragments. This has led to developments in the instrumentation, such as improvement of the direct transfer of high-molecular-weight and polar products to the ion source of the mass spectrometer the measurement of these compounds over extended mass ranges, and the use of soft ionization conditions, such as field ionization (FI) and chemical ionization (CI) instead of electron impact (EI). Methods such as FI and CI are useful due to the difficulties arising from EI, such as the variation of fragmentation depending on instrumental conditions and the fact that only low-mass ions are observed. However, one problem with the soft ionization methods is the higher cost of instrumentation (Migallon and Marín, 2001; Wolf *et al.*, 1980).

5.2 Instrumentation of pyrolysis GC/MS

A typical pyrolysis GC/MS instrument for the use in polymer analysis consists of a pyrolyzer coupled to a gas chromatograph and a mass spectrometer (Figure 5). A pyrolyzer, through which the carrier gas flows (usually He or N_2), is directly coupled to a gas chromatograph with a high-resolution capillary column.

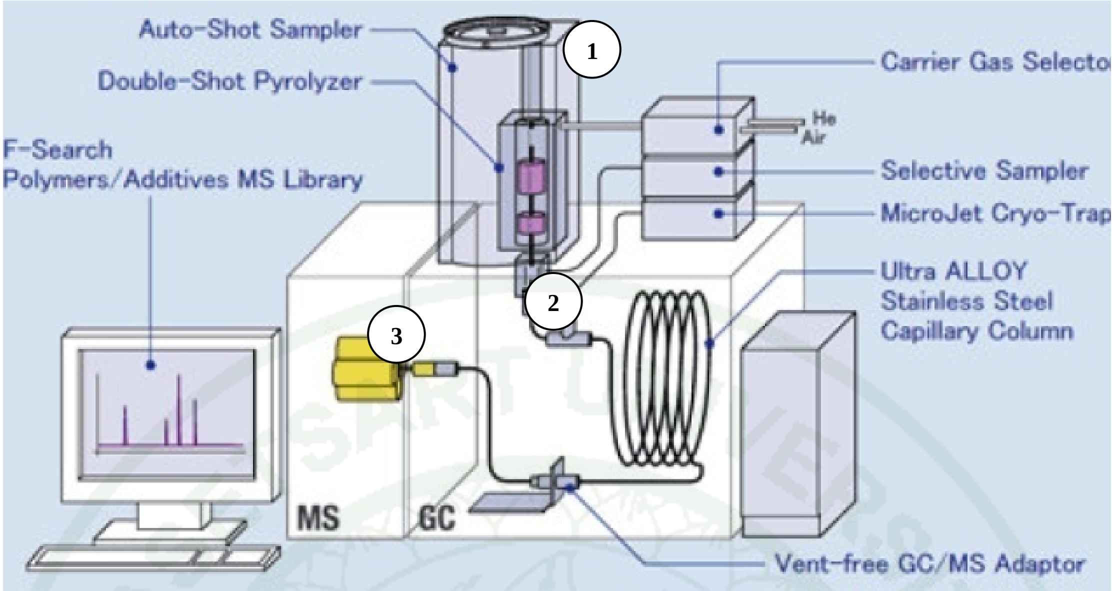


Figure 5 Diagram of typical pyrolysis-gas chromatography/mass spectrometry

Source: NSW Department of Primary Industries (n.d.)

5.2.1 Pyrolyzer

The most commonly used pyrolyzers can be classified into three groups: (1) resistively heated electrical filaments, (2) radio-frequency introduction heated wires (Curie-point pyrolyzers), and (3) microfurnace type.

Each of these designs has its particular characteristics. Thus, the filament-type permits multistep pyrolysis, which enables the discriminative analysis of formulations in a compounded material. The Curie-point type offers the most precisely controlled equilibrium temperature, but the heating conditions depend greatly on the shape of the sample holder. On the other hand, a precise temperature regulation of the microfurnace type is not easy, but it is most suited for thermally labile compounds such as biopolymers (Migallon and Marín, 2001; Wolf *et al.*, 1980).

5.2.2 Gas chromatograph

The use of highly effective capillary columns is essential for high-resolution pyrolysis GC/MS, because pyrolysis products of polymers are generally very complex mixtures. The bonded-phase fused-silica columns are especially effective for pyrolysis GC/MS because of their low level of stationary-phase bleed at elevated temperatures (Migallon and Marín, 2001; Wolf *et al.*, 1980).

The split mode is usually favored for pyrolysis GC/MS, because the very low velocity of the carrier gas for the splitless mode often causes undesirable secondary reactions of the degradation products in the pyrolyzer. However, in the split mode, the product composition entering the capillary column sometimes differs from the original, depending on the volatility of each component and the split temperature. This can cause problems in reproducibility and precision of results because the degradation products of polymer usually consist of complex mixtures with a broad range of volatilities. In order to overcome this problem, the splitter is modified and the dead volume is packed with as ordinary chromatographic packing material. The splitter is then maintained, independently, at the maximum temperature of the column. This arrangement permits the obtaining of reproducible results and protection of the instrument (Migallon and Marín, 2001; Wolf *et al.*, 1980).

5.2.3 Mass spectrometer

The advances in the specific determinations for peaks by GC/MS were significant for the development of the technique. In particular, a computer-controlled GC/MS system readily leads to the rapid and accurate identification of peaks (Migallon and Marín, 2001; Wolf *et al.*, 1980).

5.3 Methods of analysis by pyrolysis GC/MS

A very important aspect to be controlled in pyrolysis GC/MS is the reproducibility of results. This is important not only for determination of structure but also for reactions between fragments in the GC. In order to obtain high reproducibility, two different approaches can be used for a particular sample:

5.3.1 Isothermal or flash pyrolysis

The temperature of the sample is suddenly increased (10-100 ms) to reach the thermal decomposition level (500-800°C). This process can be carried out by means of a platinum-rhodium filament heated by an electrical current directly coupled to the injector port of the GC. Some pyrolysis fragments are obtained in a very short time and can be directed sent to the column and detector. In spite of this short time for pyrolysis, it is possible to indicate three different phases: (a) heating (10^{-2} - 10^{-1} s), (b) stabilization of the maximum temperature, and (c) cooling. However, the main drawback of this technique is the lack of equilibrium between temperatures with the pyrolyzer (Migallon and Marín, 2001; Wolf *et al.*, 1980).

5.3.2 Programmed temperature pyrolysis

The sample is heated at a given heating rate in a manner similar to the linear heating used in thermogravimetry or differential scanning calorimetry. The desirable conditions of this method are their uniform heating, a minimum dead volume and a device to introduce samples. Pyrolysis fragments are then analysed as a function of temperature (Migallon and Marín, 2001; Wolf *et al.*, 1980).

5.4 Applications of analytical pyrolysis in lignin analysis

Analytical pyrolysis is a powerful technique that is recently used as a new and reliable technique to study the chemistry of wood and pulps. The technique has advantages over the other analytical techniques in that little sample preparation is required and enough data can obtained from very small amounts such as dirt spots on paper. In many case good data can be obtained by analyzing the pyrolysis products by gas chromatography with frame ionization detector (FID). Detection with mass spectrometry (MS) offers the advantage of absolute qualitative and quantitative determination of the chemical matters. In many instances the chromatography can be improved by on-line derivatization of the pyrolysis products with an alkylating agent e.g. tetramethylammonium hydroxide (TMAH) as thermochemolysis (Sitholé, 2000).

5.4.1 Lignin structure

Ohra-aho *et al.* (2000) examined the behavior of lignin structural units during pyrolysis GC/MS analysis using model compounds in investigating the correlation between lignin structure and the spectrum of obtained degradation products. They were found as expected that β-*O*-4 linkages were easily cleaved, but the fate of the 5-5 linkage in biphenyl structures was found to depend on the side chain structure.

Del Río *et al.* (2007a) investigated the presence of 5-hydroxyguaiacyl moieties in the lignin from different woody and non-woody angiosperms plants using pyrolysis GC/MS. They reported that 5-hydroxyguaiacyl moiety is also involved in dehydrogenative polymerization during lignin biosynthesis. By the pyrolysis result, they are supposed to arise from the pyrolysis of 5-hydroxyguaiacyl moieties, which are supposed to be native constituents of lignin in plants forming benzodioxane as new substructure in the lignin.

Direct analysis of lignin from softwood kraft pulp by pyrolysis GC/MS was reported by Ohra-aho *et al.* (2005). They pointed out that their method was sensitive enough to detect aromatics from fully bleached softwood pulps at a concentration level of 0.4 wt.%. They studied the effect of bleaching on lignin structures in softwood pulps by following the changes in guaiacyl-type degradation product distribution. They found the residue lignin structures that has been modified during cooking were removed during the course of bleaching and they also resulted that the residual lignin in fully bleached pulps therefore was found to bear features characteristic of native lignin in addition to increased oxidation. Moreover, they observed a striking enrichment of hydroxyphenyl-type aromatic pyrolysis products during bleaching and also suggested that they may be derived not only from lignin but also from other pulp components.

Analytical pyrolysis study of mediated laccase biodelignification (laccase mediator system: LMS) of *E. globulus* kraft pulp was studied by Oudia *et al.* (2007). They used the original milled wood sample and the kraft pulps to characterize by pyrolysis GC/MS. These lignin products can be detected with high sensitivity using the selected ion chromatograms even in the bleached pulp of low lignin content (about 0.5%). They resulted that the pyrolysis GC/MS shown the lignin moieties were similarly altered during biobleaching as during pulping, which is exemplified by the preferential removal of aldehyde groups from the alkyl side groups.

5.4.2 Lignin content

Ohra-aho *et al.* (2000) pointed out another goal of their research that they used an analytical pyrolysis to develop method that would give not just qualitative but also quantitative information. This was accomplished by using a weighed amount of sample and normalizing the combined peak area of the degradation products to the weight of sample. Thus, their method can be used to obtain information about the proportion of aromatic lignin in isolated lignin samples or about the content of aromatic lignin in the fiber.

Sonoda *et al.* (2001) proposed a new method to quantify lignin monomer composition in detail by pyrolysis gas chromatography combined with preliminary acetylation of lignin samples. They used dehydrogenative polymer (DHP) as lignin model compounds to determine the lignin monomer compositions. The lignin acetylation procedure would contribute to prevent secondary formation of cinnamaldehydes from the corresponding alcohol forms during pyrolysis, which are otherwise unavoidable in conventional pyrolysis GC process to some extent. They concluded that the new pyrolysis GC method combined with sample preacetylation allowed an accurate quantitative analysis of detailed lignin monomer composition using a microgram order of extractive-free plant samples.

Many research groups studied the quantitative analysis of lignin in wood by pyrolysis GC/MS. Alves *et al*. (2006b) investigated the quantification of the lignin content (pyrolysis lignin) of Maritime pine and spruce wood samples directly from the programs. They compared the precision between the pyrolysis lignin method and reference Klason method and found that it good correlation between pyrolysis lignin and Klason lignin content for both species which allows to calculate the Klason lignin content from pyrolysis lignin content for both species using only one model. Later, Alves *et al*. (2008), part 2 of their work, studied the evaluation of the common pyrolysis lignin model based on pine and spruce with larch wood samples. Their result was found that only small difference between the measured and the predicted Klason lignin contents of larch and the larch samples could be include in the model. Moreover, they also investigated the influence of compression wood and the result revealed small difference between normal wood and compression wood. Therefore, they concluded that this so-called softwood model can be used for pine, larch, and spruce wood with the limitation of the highest and lowest values and they expected that this model can also be used to predict Klason lignin contents of other softwood species for both normal wood and compression wood.

5.4.3 S : G : H ratio of lignin

Comparisons of lignin S/G ratio by various methods were reviewed by Yokoi *et al*. (1999). They summarized that utilizing chemical degradation, such as permanganate- and alkaline nitrobenzene oxidations, followed by chromatography have mainly been employed for determination of lignin S/G ratio in woods. However, these methods required not only tedious and time-consuming but also fairly large amounts of wood sample, up to 100 mg). Furthermore, often causes denaturation of lignin structures by strong acidity of reagents used in degradation process obtained the results that are not reflect the original structure of lignin. In thioacidolysis method, in which lignin is solvolyzed with dioxane and ethanethiol, under relatively mild acidic conditions is expected to yield more representative results for the original chemical structure of lignin. However, is usually takes several hours for the thioacidolysis reaction and derivatization for GC analysis. The spectroscopic analyses, such as FT-

IR and NMR spectroscopy, have also been utilized for qualitative and/or quantitative analysis of the functional groups in lignin. However, they are difficult to determine the precise S/G ratio in lignin owing to both insufficient sensitivity and poor resolution of the spectra. A rapid and highly sensitivity method for characterizing the structure of lignin was found in the pyrolysis GC method. This technique can be classified the tree species based on the distribution of lignin-derived pyrolyzates in trace wood samples. Lima *et al*. (2008) studied the comparison between analytical pyrolysis and nitrobenzene oxidation for determination of S/G ratio in Eucalyptus *spp*. lignin. A total of 41 compounds were identified using pyrolysis GC/MS, of which 11 were derived from carbohydrates and 30 from lignins. They compared the calculated S/G ratio on the basis of the areas of peaks recorded in the programs with the results obtain by alkaline nitrobenzene oxidation. They concluded that the selected lignin markers were efficient for the determination of S/G ratio in eucalyptus wood by pyrolysis GC/MS. Therefore, pyrolysis GC/MS technique produced results that are comparable to the nitrobenzene oxidation method, with the advantage of required small wood samples and a short analysis time.

Rodrigues *et al*. (1999) determined the tree-to-tree variation in S/G ratio of *E. globulus* wood lignin by analytical pyrolysis (pyrolysis GC/MS). They concluded that the analytical pyrolysis method proved reproducible and sensitive for the measurement of the natural variability of *E. globulus* with regard to lignin S/G ratio and also pointed out that the pyrolytic determination of the S/G ratios is recommended as an evaluation trait in breeding programmes for pulpwood production. The application of discriminative analysis of *E. camaldulensis* grown from seeds of various origins based on lignin S/G ratio measured by pyrolysis GC using vertical microfurnace pyrolyzer was evaluated by Yokoi *et al*. (2001). On the basis of S/G ratio calculated from molar yields of peak specifically derived from pyrolyzates characteristic to lignin was successfully used to discriminative analysis among the eucalyptus trees using principal component analysis (PCA).

Del Río *et al*. (2005) determined the influence of eucalyptus lignin composition in paper pulp yield using pyrolysis GC/MS. The integrity of the lignin/carbohydrate ratio was assessed through comparison to lignin S/G ratio. Their result was found that a significant correlation was observed between pulp yield and S/G ratio, but the lignin/carbohydrate ratio did not show a significant correlation. This indicates that the S/G ratio is a more important parameter influencing pulp yield than the lignin content.

The use of analytical pyrolysis data to develop an NIR method for determining H/G ratios, with a focus on assessment of the natural variation of lignin composition in maritime pine wood was evaluated by Alves *et al*. (2006a). They concluded that NIR spectroscopy can be used to assess the lignin composition (H/G ratio) of maritime pine wood and its precision is comparable to that of the reference pyrolysis GC/MS method.

Greenwood *et al*. (2002) studied the potential for using lasers as a highly directional thermal source in analytical pyrolysis. They investigated by the laser micropyrolysis GC/MS analysis of lignin moieties. They observed the high levels of correlation from the both laser micropyrolysis GC/MS and the traditional used pyroprobe pyrolysis GC/MS of both lignins from Douglas fir and *Quercus nigra* water oak.

There are many groups of research studied the lignin S/G ratio by pyrolysis in the presence of tetramethylammonium hydroxide (pyrolysis/TMAH) (thermochemolysis or simultaneous pyrolysis methylation). Because of the specificity of the pyrolysis GC/MS method is low due to it randomly destroys the subunits in lignin that provided less structural information. For the structural elucidation of the polymer, pyrolysis/TMAH is a novel technique that has provided detailed information on both the structure and the composition of polymers. Kuroda *et al*. (2002) explored the feasibility of TMAH/pyrolysis GC/MS as a practical technique for characterizing lignin. Their research presented the determination of S/G ratio by pyrolysis in the presence of TMAH agreed well with determined by thioacidolysis with a significant linear regression ($R^2 = 0.9867$). Thus, their result showed that pyrolysis with TMAH is an effective tool for obtaining information on the S/G ratio for in situ lignin. Again, González-Vial *et al*. (1999) used flash-pyrolysis in the presence of TMAH of woods from different species of eucalyptus to yield series of G and S units. Their pyrolytic data were in agreement with those obtained from CP/MAS ^{13}C-NMR spectral data of the same woods by showed additional differences in pyrolytic patterns, which might be used as a basis for chemotaxonomic differentiation of the origin of the woods. Latest, Del Río *et al*. (2007b) used the pyrolysis/TMAH-GC/MS to analysis the composition of non-woody plant lignins and cinnamic acids. They mentioned that *p*-hydroxycinnamic acids, *p*-coumaric and ferulic acids, as well as different diferulates, are abundant in non-woody plants forming cross-linkages between lignin and polysaccharides that have been especially investigated in grasses. The presence of these phenylpropanoid compounds constitutes a complication for lignin analysis by analytical pyrolysis since they yield products similar to those of corresponding lignin units. This problem can be solving by using pyrolysis in the present of TMAH, that releases intact *p*-hydroxycinnamic acids as methylated derivatives.

6. NIR Spectroscopy

6.1 Introduction to principle of NIR spectroscopy

Near infrared (NIR) spectroscopy, as defined by IUPAC, is based on the absorption of electromagnetic radiation at wavelengths in the range of 780-2500 nm corresponding to the wave number range 12820-4000 cm^{-1}. It covers the wavelength range adjacent to the mid infrared and extends up to the visible region (Figure 6). The major features that seen in the NIR region result from overtones of the fundamental bond stretching or bending vibrations and combinations. These arise because of anhamonicity in the potential energy well. The overtones are due to vibrational transitions to levels above the first excited state. The frequencies of these transitions are approximated by integral multiples of the fundamental frequency. Combinations occur when light is absorbed by more than one bond attached to a common atom. Every molecule containing hydrogen will have a measurable NIR spectrum and given the ubiquitous distribution of hydrogen this means that a very large number of analytes are amenable to NIR analysis. The formation of hydrogen bonds changes bond lengths and this changes the bond strength. The result of this is an apparent shift in the expected positions of absorption peaks (Davies, 1998; Pope, 1995; Osborne, 2000).

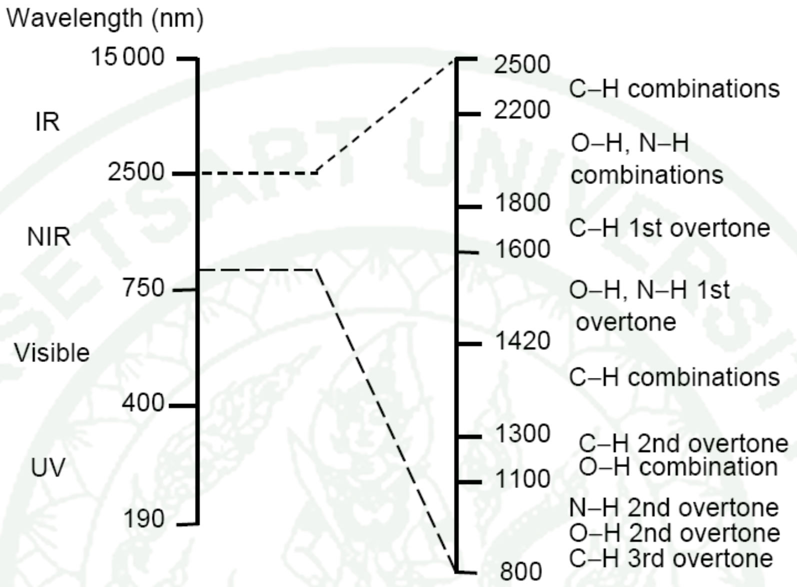


Figure 6 Principal types of NIR absorption bands and their locations

Source: Osborne (2000)

6.2 Instrument and measurement for NIR spectroscopy

A NIR spectrometer is generally composed of a light source, a monochromator, a sample holder or sample presentation interface, and a detector, allowing for transmittance or reflectance measurements (Reich, 2005). Approaching to NIR measurement, light comes from a source and illuminates the sample. The light interacts with the sample and modified in specific way. The modified reflected or transmitted light is directed to a detector that is sensitive to NIR light. The signal is converted into electrical information that can be read by a computer. This is generating spectral data are in the computer applied to the calibration models for quantitative and qualitative analyses. Figure 7 shows the principle of NIR measurement.

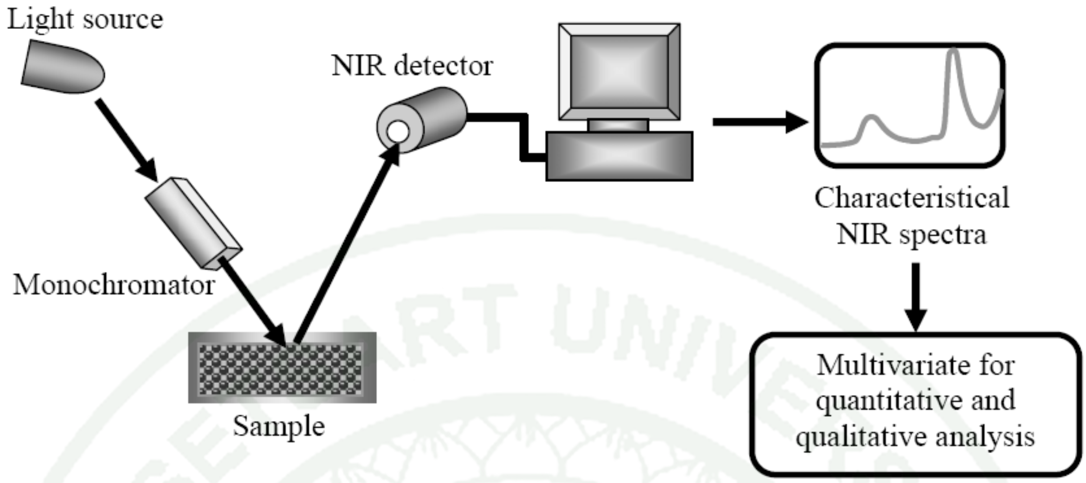


Figure 7 Principle of NIR measurement

Source: Ritthiruangdej (2006)

The quartz-halogen lamp is by far the most popular source of NIR energy. Quartz-halogen lamps are quite popular now because of their widespread use for household lighting. Light-emitting diodes are the second most popular source of NIR energy. Wavelength isolators range in technology; photodiode arrays (PDA), diode array detector (DAD), laser diodes (LD), fixed filters (FF), wedge interference filters (WIF), tilting filters (TF), acousto-optical tunable filter (ATOF), liquid crystal tunable filters (LCTF), prisms, grating, Fourier transform NIR (FT-NIR), and Hadamard. Spectrometer designers have a number of detectors from which to choose, the most popular of which are lead sulfide (PbS)-, indium gallium arsenide (InGAS)-, and silicon-based devices (voltaic cell, diodes, and transistors for the Herschel region). Silicon-based detectors are restricted by their spectral response to ultraviolet, visible, and short-wavelength NIR (200-1,000 nm) (Mcclure and Tsuchikawa, 2007).

Only within last 20 years has the Fourier Transform-Near Infrared (FT-NIR) instrumentation become available. Even then, the first commercial instruments had a distinct disadvantage compared to grating-based scanning instruments. FT-NIR spectrometers employ as entirely different method for producing spectra. There is no dispersion involved. Energy patterns set up by an interaction with a sample and a reference and moving mirrors (or other optical components) produce sample and reference interferograms that are used to calculate the absorbance spectrum of the sample. There are two advantages of FT-NIR spectrometers that make it attractive. The first is the throughput advantages. In the absence of dispersion, the energy at the output of an FT-NIR interferometer (similar to a monochromator) can be many times greater than that obtained from a grating monochromator. Nonetheless, the first commercial instruments literally starved for output energy, resulting in noise levels incommensurate with the standards set by dispersion-type spectrometers. More recently, manufacturers have improved the noise levels of FT-NIR spectrometers to the point where they readily compete with all other types of NIR technology (Mcclure and Tsuchikawa, 2007).

The appropriate NIR measuring mode is dictated by the optical properties of the samples. Figure 8 illustrates some of measuring modes known as transmission, reflection, transflection and interaction.

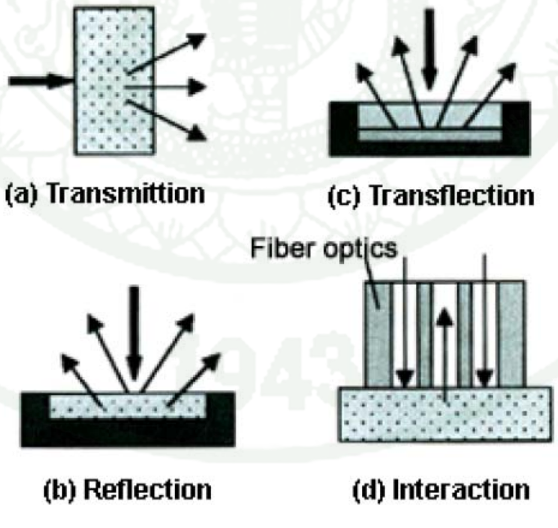


Figure 8 Type of sample presentation

Source: Mcclure and Tsuchikawa (2007)

Sample selection, sampling, sample preparation and sample presentation to the instrument are fundamental to accurate and precise testing by both NIR and reference methods. Most analysis is carried out on what is purported to be a sample that is representative of the whole population and provides information on what the operator needs to know. This can be achieved only by continuous sampling or repeated random sampling of the materials. Sample presentation includes documentation, blending, sub sampling, removal of unwanted material, grinding and storage. Sample presentation is also the important factor affecting NIR measurements. The definition of sample presentation is meant how to present or set a sample to NIR instrument. Therefore, several types of sample cells, such as quartz cuvettes with defined optical path length for liquids, specifically designed sample cells with quartz windows for semi-solids and powders, and adjusted sample holders for tablets and capsules have been developed by the manufacturers and the users of NIR instruments.

6.3 General procedures for development of NIR analysis

Typical steps in developing a NIR analysis procedure include optimizing the instrument and sampling system, establishing a background or reference spectra, validating the primary analysis method, collecting spectra from the calibration set, processing the spectral data to improve resolution and removing sampling artifacts, applying a modeling technique to the data (usually a linear model such as multiple linear regression (MLR), principal component regression (PCR), or partial least squares regression (PLSR)), and validating the technique by predicting analytical values for sample independent of the calibration set. Statistical tests may be applied to detect outliers both in the calibration and validation sets (Pope, 1995).

Generally, NIR calibration sets range from 20-400 samples. The number of samples required for robust calibration will vary with the nature of the calibration set, the number of independent variables, and the precision desired. After confidence in particular model has been developed through the use of validation samples, the models might be compared over a period of time to determine which is the most robust with respect to environmental factors such as temperature, humidity and process variation (Pope, 1995).

6.4 Basic of NIR calibration

For both of qualitative and quantitative analyses, NIR spectroscopy needs a calibration equation. The calibration procedure involves collecting a number of samples, obtaining both reference and NIR data on each sample and deriving a calibration equation from these data by using chemometrics. The calibration process basically involves the following steps: 1) selection of a representative calibration sample set; 2) Determination of standard concentration and NIR spectra acquisition; 3) Development of mathematical model (multivariate analysis) by relates the spectral variations to the reference values of the analytical target property; 4) Validation of the calibration model by cross validation, set validation or external validation. Figure 9 shows the calibration and validation process.

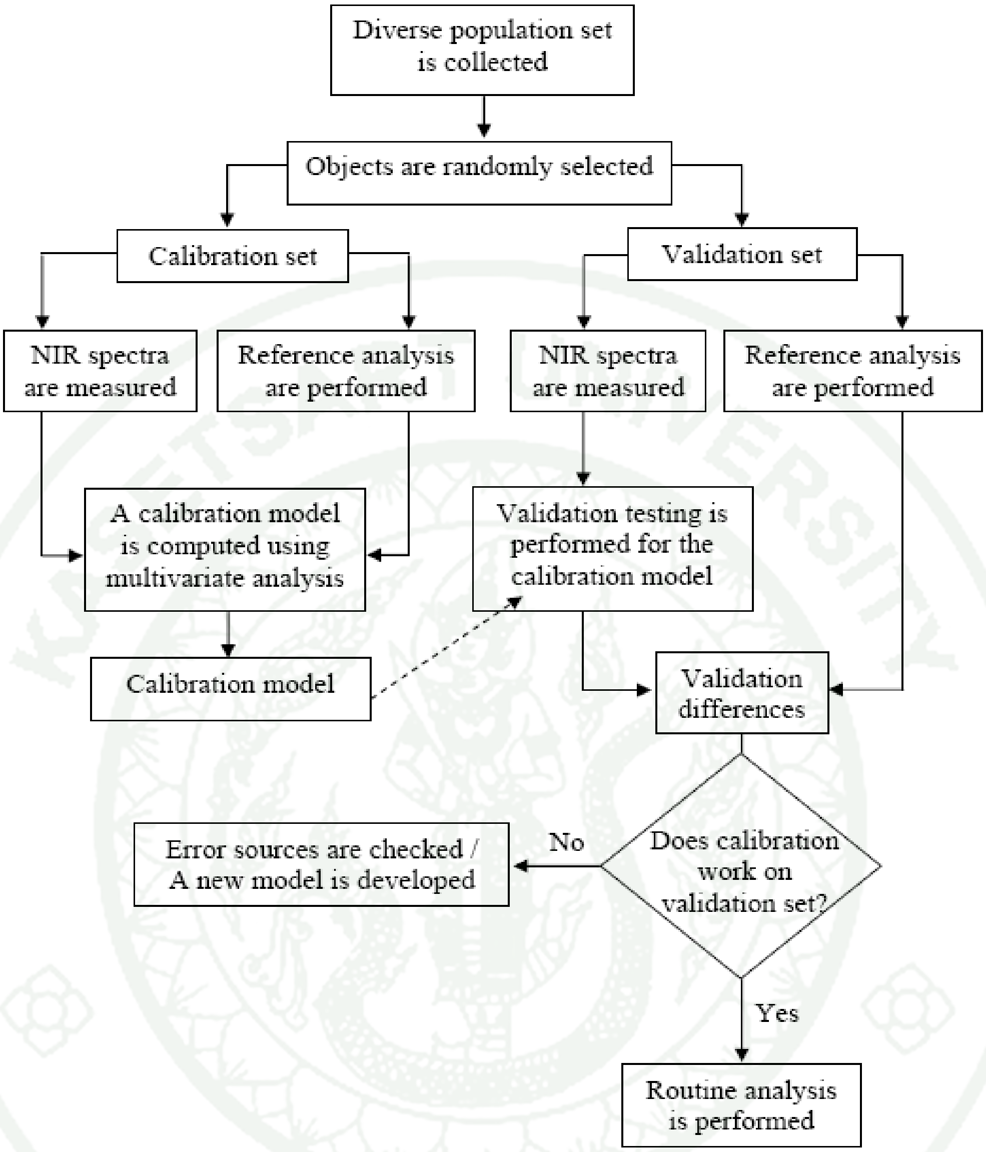


Figure 9 Flow diagram of NIR calibration and validation process

Source: Ritthiruangdej (2006)

6.5 Advantages of NIR analysis

NIR spectroscopy has many advantages in analysis of wood and pulp chemistry when compared to chemical and other instrumental methods, although the results from the other techniques are more accuracy than those from the NIR spectroscopic method. The advantages of NIR spectroscopy superior the other methods are summarized by following:

- Non-destructive and environmental friendly method
- Rapid speed and reliability of determinations
- Rapid and easy sample preparation
- small amount of samples and/or chemicals
- Low cost analysis in long term consideration
- Multiplicity of analysis with one operation
- Operation by unskilled personal

From the many previous literature reviews can be concluded that the NIR spectroscopy could be substituted the conventional method in the routine analysis of pulp and paper field. Thus, the pulp and paper industry in Thailand should be rapid development for the advancement, safety for environment, and reduce cost for the future business competition.

6.6 Applications of NIR assessment for pulpwood properties

The work of Karl Norris of the USDA was central to the development of NIR as a tool for analysis of agricultural products. By the early 1990s the analytical applications of NIR spectroscopy were widespread with many different industries (agriculture, food, paper, petro-chemical, polymer and textile) utilizing the technology (Pope, 1995; Schimleck, n.d.).

Nowadays studies concentrated on the feasibility of NIR spectroscopy to predict wood chemistry, specifically pulp yield, and two wood components (cellulose and lignin) that together have a large influence on pulp yield. A long-term objective of tree breeding programs has been to improve the pulp yield of plantation grown trees as pulp yield is an important factor in determining plantation profitability. Thus NIR spectroscopy was a rapid and non-destructive method (avoid to cut down the trees for sampling) had long been sought for this purpose (Schimleck, n.d.; Baillères *et al.*, 2002; Kelley *et al.*, 2004; Jones *et al.*, 2006, 2008).

The majority of work with NIR has focused on chemical composition of wood properties. Several publications have highlighted the use of NIR spectroscopic techniques to predict the wood lignin content and its composition. Poke *et al.* (2004), Poke and Raymond (2006), Hodge and Woodbridge (2004), Baillères *et al.* (2002) and Jones *et al.* (2008) used diffuse reflectance NIR spectra to characterized the lignin with mostly use long-NIR wavelength between 1100-2500 nm, whereas Yen *et al.* (2004) and Yamada (2006) used transmittance NIR spectra with a range of 600-1900 nm to characterized the lignins. Moreover, Kelley *et al.* (2004) collected the reflectance NIR spectra using interaction mode via a fiber optic probe with the full-spectral range of NIR combined to visible region (400-2500 nm) to characterize the lignin. All of their research results obtained a strong correlation between the NIR results and those obtained from conventional wet chemistry results for the lignin content and S/G ratio. Their reported predictions were very good.

In addition to the use of NIR for chemical properties of wood, many researches are paying attention to the applicability of NIR spectroscopy with the aid of multivariate analysis as chemometrics to evaluate various physical properties of wood, such as moisture content, grain angle, surface roughness, density, anatomical parameter and also mechanical properties of wood etc. Schimleck and Evans (2003) investigated the air-dry density calibration of radiate pine wood using five factors giving an excellent relationship between SilvisScan-2, which utilized X-ray diffraction, X-ray densitometry, and image analysis to measure a range of wood properties from on increment core, determined air-dry density, and NIR fitted air-dry density. The study of their work indicated that NIR spectroscopy provided a rapid method for the routine analysis of air-dry density in large numbers of core samples. Via *et al.* (2005) studied the ability of NIR spectroscopy to monitor air-dry density distribution and variation of wood. Their work indicated that the use of NIR was successful at roughly estimating the variation. The study of Schimleck and Evans (2004) with *Pinus radiata* D. Don utilized data provided by the SilviScan instruments to develop NIR calibrations for tracheid coarseness and also wall thickness. Their work obtained the coefficient of determination (R^2) 0.91 for coarseness model. Morphological characteristics of tracheid, including coarseness, specific surface, wall thickness, perimeter and radial and tangential diameter of *Pinus taeda* L. are investigated by Jones *et al.* (2005). They found that the properties related to density are strong correlations, particularly the strongest R^2 being 0.80 for coarseness and 0.84 for wall thickness. Thus, according for the results of many researchers, it may be concluded that the density related properties can be estimated by NIR with sufficient accuracy to be used in operational setting.

Recently, the use of NIR to measure wood properties of interest to pulp and paper industry has received a great deal of attention. Pulp yield is a most important factor for the pulp and papermaking industry. Wright *et al.* (1990) reported the applicability of NIR reflectance spectroscopy to predict the pulp yield from wood at an early stage of NIR research related to wood. Michell and Schimleck (1998) and Schimleck *et al.* (2005b) reported the feasibility of the rapid assessment of pulp yield of clonal eucalyptus and hybrid poplar plantations by NIR spectroscopy. Correlation between pulp yield predicted for increment cores and whole-tree value indicated that useful predictions of whole-tree pulp yields could be obtained from measurements on increment cores (at 10% tree height). Schimleck and French (2002) reported the application of NIR spectroscopy to clonal *E. globulus* samples covering a narrow range of pulp yield. Furthermore, Schimleck *et al*. (2004, 2005a) introduced some research concerning with kraft pulp yield in *E. nitens* wood, where they paid attention to genetic gains based on increment core NIR spectra. Ruiz *et al*. (2005) evaluated the estimating pulp yield of *E. globulus* wood using NIR spectroscopy comparison with mid-infrared (MIR) spectroscopy. They reported that more precise results were obtained based on NIR spectra where the calibrations models were more robust and better predictors than obtained for the MIR spectra. Within-tree variation in kraft pulp yield, predicting using NIR reflectance analysis, was investigated by Raymond *et al*. (2001) to develop a non-destructive sampling strategy. They sampling simulated core samples at six fixed heights easily accessible from the ground (0.5, 0.7, … 1.5 m) and at seven percentage heights (0, 20, 30, … 70%). Whole-tree values, calculated from percentage height data, were correlated with the core data to determine the optimal sampling height. They found that core samples for *E. globulus* recommended sampling height at 1.1 m were found to be good predictors of whole-tree pulp yield, whereas results for *E. nitens* on high quality sites recommended sampling height at 0.9 m were found to be good predictors of whole-tree pulp yield.

In Thailand, the pulp and paper industries have played an important role and their productions are expanding significantly year by year. Especially, *E. camaldulensis* is used as the raw material for pulp and papermaking in Thailand. Terdwongworakul *et al*. (2005) investigated the feasibility of predicting chemical properties and pulp yield using NIR spectroscopy of *E. camuldelensis* in Thailand tree plantations for improving wood selection for high quality pulp. They concluded and expected their NIR developed models that NIR spectroscopy could be satisfactorily used as an effective assessment technique for tree breeding and clone selection programs for *E. camuldelensis* in Thailand.

7. Chemometrics for NIR

Only NIR information cannot be used to determine analyse concentrations directly because of the way in which near infrared radiation passes into, and reflect from the sample. Thus, chemometrics are necessary to make the calibration model.

7.1 Introduction to chemometrics

Chemometrics is the discipline using computers and mathematics to derive meaningful chemical information from samples of varying complexity. A main part of chemometrics is multivariate data analysis, which is essential for both quantitative and qualitative assays based on NIR spectroscopy. Quantitative analysis of NIR spectroscopic data is based on Beer’s law, which states that a linear relationship exists between the molar concentrations of a substance at a given wavelength.

The multivariate regression methods most frequency used in quantitative NIR analysis are principal component regression (PCR) and partial least squares regression (PLSR). Both PCR and PLSR are full-spectrum calibration methods which achieve the same end by reducing the amount of spectral data in another way, constructing a small number of factors without discarding any useful information. PCR uses the principal components provided by principal component analysis (PCA) to perform regression on the sample property to be predicted. PLSR finds the directions of greatest variability by comparing both spectral and target property information with the new axes, called PLS components or PLS factors or latent variables (Reich, 2005).

7.2 Pretreatment techniques for NIR spectra

NIR bands are broader than those in mid-infrared and the spectra quite often suffer from the problems of unwanted spectral variations and baseline shifts. These may have some sources such as light scattering from solid samples or cloudy liquids, poor reproducibility of NIR spectra caused by pathlength variation, variation in temperature, packing density and size of samples, and the various kinds of noise such as from detector or AC to DC converter. Thus, NIR information usually uses the data pretreatment techniques before development of NIR calibration model. There are two most favorite techniques uses namely derivative and multiplicative scatter correction (Ozaki *et al*., 2007).

7.2.1 Derivative

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.

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 10.

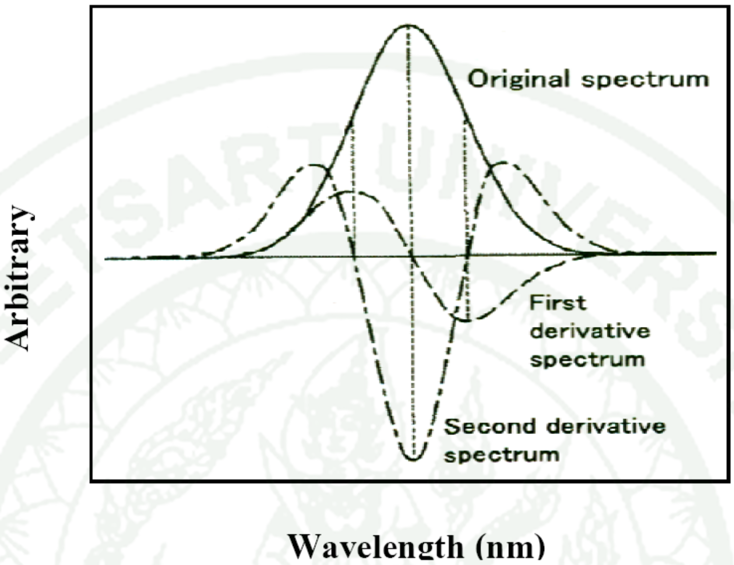


Figure 10 First- and second-derivatives of the original spectrum

Source: Weantong (2005)

7.2.2 Multiplicative scatter correction (MSC)

MSC is a powerful method for correcting vertical variations of the baseline (additive baseline variation) and inclination of the baseline (multiplicative baseline variation). The basis of MSC lies in the fact that light scattering has wavelength dependence different from that of chemically based light absorption and light scattering. Thus we can use data from many wavelengths to distinguish between light absorption and light scattering (Ozaki *et al*., 2007).

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 11(A), it is apparent that the largest source of variance within the 40 samples is derived from scattering. Figure 11(B) presents the plots of the spectra are pretreated with MSC. The spectra demonstrate the potential of MSC improves essentially the linearity in NIR spectroscopy.

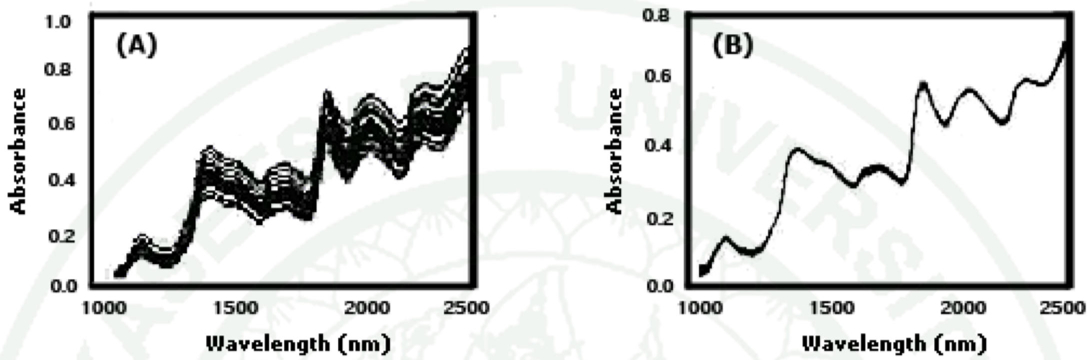


Figure 11 Forty corn calibration samples: (A) untreated and (B) MSC applied to 40 calibration samples

Source: Suttiwijitpukdee (2006)

7.3 Multivariate data analysis for quantitative NIR calibration development

To develop a prediction equation for quantitative chemical analysis, multiple variable analysis routines are run on a calibration set of spectra with known concentrations. The multivariate data analysis is often used in spectroscopy to extract information from complex spectra containing overlapping absorption peaks, interference effects, and instrumental artifacts from the collected data. These techniques are usually based on linear regression algorithms. The change in instrument response is correlated with independently determined primary values to develop predictive models. In multiple linear regression (MLR), changes in absorbance intensity at several wavelengths are correlated with changing analytical values. In simple chemical systems, it is expected that will be required for each independently varying component of the chemical matrix. Other methods of linear regression analysis, termed chemometrics, may be applied to more complex systems.

Among these are included principal component regression (PCR) and partial least squares (PLS). These algorithms calibrate using all variable spectral data. Both PLS and PCR represent the original data as linear combinations of factors weighted by scores. In general, simple chemical systems, such as mixtures, are well represented with simple models, such as MLR. Complex systems, such as biological materials, usually require complex model, such as PLS or PCR. Brief explanation of MLR, PCA and PLS are following (Pope, 1995):

7.3.1 Multiple linear regression

The purpose of multiple regressions is to find the relationship between several independent or predictor variables and a dependent or criteria variable. Multiple linear regression (MLR), or called inverse least squares, is a quantitative calibration model that use the selective wavelength method to establish a relation between NIR spectra and quantifiable sample property, e.g. contents of components such as lignin, cellulose, pulp yield, etc.

The basis for an MLR calibration is two sets of data: (1) the set of value measured at the different wavelengths for the calibration set samples and (2) the set of 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 squares of the vertical distances of the property values from the line. The regression line is calculated by using the following equation:

$$y = F_0 + F_x$$

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 Beer's law applies, e.g. in case of UV/VIS spectra of diluted solution. For NIR analysis of solid or liquid samples, however, there will always by 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:

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

However, selective wavelength method likes MLR may have some problems i.e. underestimation and/or over fitting because of the excessive independent variables. Therefore the full-spectrum analysis methods, PCA and PLS regressions, are the way that nowadays useful to solve these problems. PCA and PLS analysis are explained in the 7.3.2 and 7.3.3 sections.

7.3.2 Principal component analysis

Principal component analysis (PCA) is one of the variable-directed techniques for forming new variables which are linear composites of the original variables. The objective of the PCA is to reduce the dimensionality of the data set. It involves a mathematical procedure that transforms a number of possibly correlated variables into a smaller number of uncorrelated variables that are called principal components (PCs). The values of new variables are called principal components scores. The PCA principle is to find the directions in space along which the distance between data points is the largest. This can be translated as finding the linear combinations of the initial variable that contribute most to make the sample different from each other. The second PC accounts for the maximum variance that has not been accounted for by the variable, and so on. Figure 12 present the model that describes the 2 new principal components variables (PC1, PC2) are derived from 3 old variables (X_1, X_2, X_3). The new axes called principal components (PCi) are derived from linear combinations of the original variables (Xi) (Lohninger, 1999).

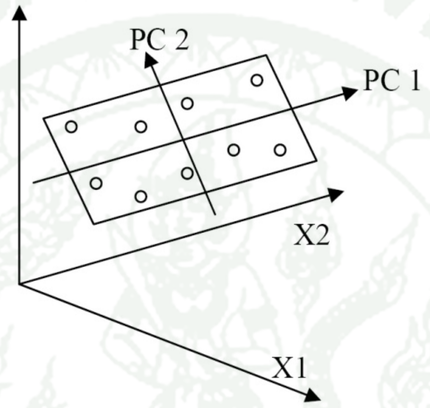


Figure 12 Principal component analysis (PCA): Reduce the dimensions of measured variables to the representative principal components (PCs)

Source: Ritthiruangdej (2006)

PCA is based on a decomposition of the data matrix X into two matrices V and U as shown in Figure 13. The two matrices V and U are orthogonal. The matrix V is usually called the loadings matrix, and the matrix U is called the scores matrix. The loadings can be understood as the weights for each original variable when calculating the principal component. The U matrix contains the original data in a rotated coordinate system (Lohninger, 1999).

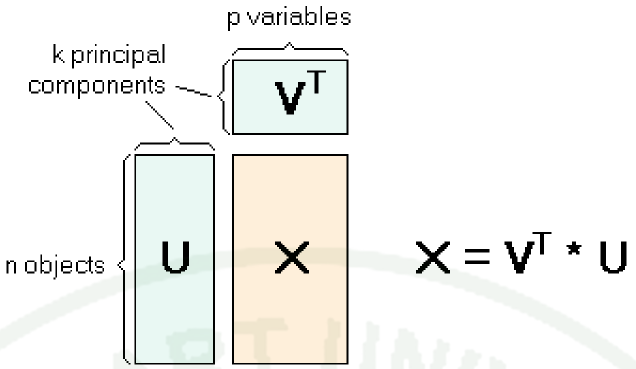


Figure 13 Scheme for explanation of principal component analysis (PCA)

Source: Lohninger (1999)

Loadings describe the data structure in terms of variable correlation. Each variable has a loading on each PC. On geometrical terms, a loading is the cosine of the angle between the variable and the current PC. A smaller angle has a larger loading. If two variables have high loading along the same PC, it means that their angle is small, which in turn means that the two variables are highly correlated.

Scores describe the data structure in terms of sample patterns, and more generally show the differences or similarities of sample. Each sample has a score on the PC. It reflects the sample location along that PC. Sample with close scores along the same PC are similar due to they have close values for the corresponding variables. Conversely, samples whose scores differ much are quite different from each other with respect to those variables.

Principal component regression (PCR) is a simple extension of MLR and PCA. In the first step, the principal components (PCs) are calculated using PCA. The scores of the most important principal components are used as the basis for the multiple linear regression with the target data y (Figure 14) (Lohninger, 1999).

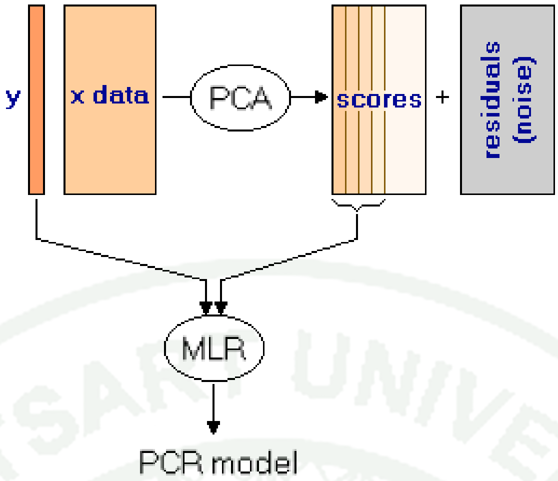


Figure 14 Steps for making principal component regression (PCR)

Source: Lohninger (1999)

7.3.3 Partial least squares regression

Partial least squares (PLS) regression is the most popular multivariate data analysis method and has been widely applied in the field of chemometrics, in sensory evaluation, and more recently, in the imaging data analysis. PLS is used as full-spectral calibration method for NIR analysis (Lohninger, 1999).

PLS models are based on principal components of both the independent data (X) and the dependent data (Y). The central idea is to calculate the principal component scores of the X and the Y data matrix and to set up a regression model between the scores (not the original data). Figure 15 describes that the matrices t and u are the principal component scores of the X and Y data. A regression set up between both scores called PLSR model (Lohninger, 1999).

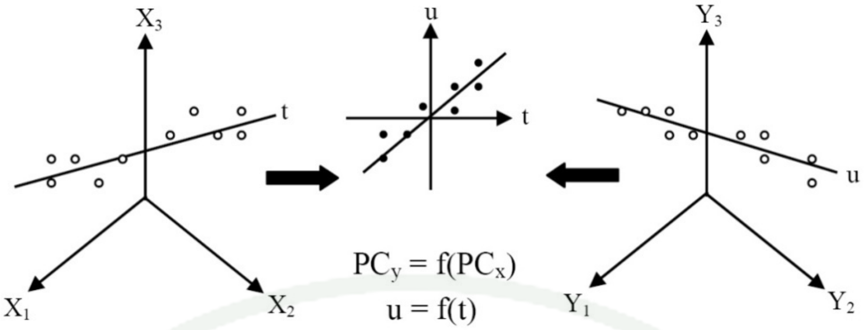


Figure 15 PLS decomposition

Source: Ritthiruangdej (2006)

Both PLS and PCR are projection methods as same as PCA. The difference between PLS and PCR is that the PLS uses both the independent and dependent variables to find the regression model, whereas the PCR decomposes the X-matrix by PCA and then fits a MLR model using the PCs instead of the original data as predictors. The scheme for explanation of PLS is shown in Figure 16.

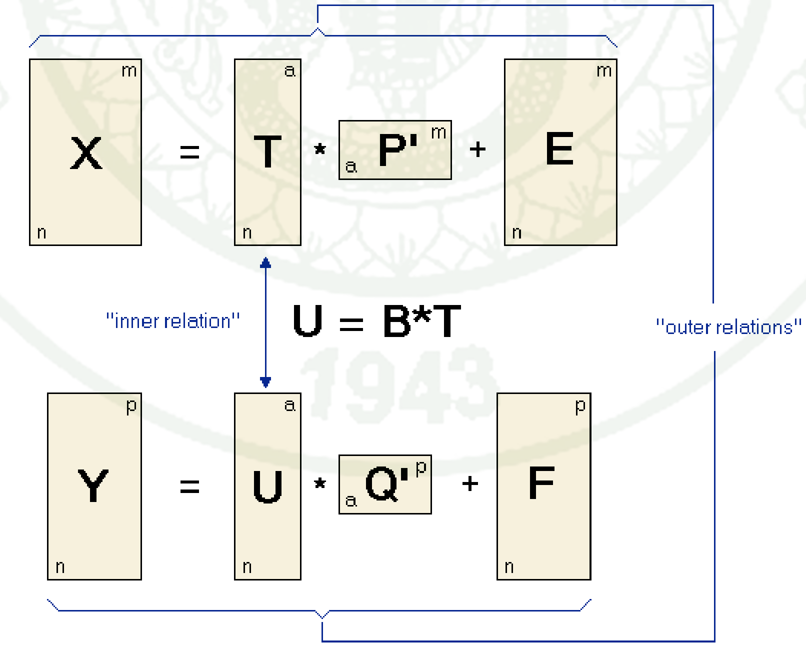


Figure 16 Scheme for explanation of partial least squares (PLS) regression

Source: Lohninger (1999)

Thus the matrix X is decomposed into a matrix T (the score matrix) and a matrix P' (the loadings matrix) plus an error matrix E. The matrix Y is decomposed into U and Q and the error term F. These two equations are called outer relations. The goal of the PLS algorithm is to minimize the norm of F while keeping the correlation between X and Y by the inner relation U = BT (Lohninger, 1999).

The important point when setting up a PLS model is to make a decision for the optimum number of principal components involved in the PLS model. While this can be done from variation criteria for other models, for PLS the optimum number of components has to be determined empirically by cross-validation of the PLS model using an increasing number of components. The model with the smallest predictive error sum of squares value can be regarded as the "best" model.

7.4 Statistical terms necessary to evaluate the calibration performance

Mathematical models that involved validation in NIR analysis are typically evaluated using coefficient of determination (R^2), bias, standard error of calibration (SEC), standard error of prediction (SEP), and ratio performance to deviation (RPD).

R^2 value is interpreted as the proportion of variance in Y data that is attributable to variance in the X data which is explained by the model. For example, an R^2 value of 0.941 means that 94.1% of the variance in Y data can be attributed to variance in X data (and 5.9% of the variance is attributable to all of the other factors). Always remember that the R^2 value is positive sign. Perfect agreement of between X and Y data will result in a R^2 value of 1.0 and in practice, this result is impossible because there are always error in both X and Y.

SEC is interpreted as the standard deviation of the true (laboratory analysis) versus the fitted (modeled) values of samples in the calibration set. It means that the SEC is used to measure goodness of fit of the model. SEP is interpreted as the standard deviation of the true versus fitted values for samples independent of the calibration set. This sample set is the prediction or validation sample set. The formulas of SEC and SEP are shown following:

$$\mathrm{SEC} = \sqrt{\frac{\sum_{i=1}^{N_C} (Y_i - \hat{Y}_i)^2}{N_C - k - 1}}$$

$\hat{Y}_i$ = the value of the constituent of interest for a validation sample *i* estimated using the calibration

Y_i = the known value of the constituent of interest of sample *i*

N_C = the number of samples used to obtain the calibration

k = the number of factors used to obtain the calibration

$$\mathrm{SEP} = \sqrt{\frac{\sum_{j=1}^{N_P} (Y_j - \hat{Y}_j)^2}{N_P - 1}}$$

$\hat{Y}_j$ = the value of the constituent of interest for a validation sample *j* predicted by the calibration

Y_j = the known value of the constituent of interest for sample *j*

N_P = the number of samples in prediction set

Ratio of performance to deviation (RPD: ratio of the SD to the SEP of the reference data used in validation), firstly defined by Phil Williams, is a measurement of the ability of an NIR model to predict a constituent. The formula of RPD is SD/SEP or $1/(1-R^2)^{1/2}$, but the former is easier to remember than the later formula. Reporting the SEP alone may be misleading unless it is reported by comparison with the SD of the original reference data. The RPD should be high if the calibration model is effective. If the SEP is close to the SD, then the NIRS calibration is not efficiency to predict the composition or functionality. If SEP = SD, the calibration is essentially predicting the population mean. An RPD below 2 cannot give a relevant prediction. An RPD value of 2.0-3.0 is regarded as adequate for rough screening. A value of above 3.0 is regarded as satisfactory for screening that recommended for NIRS applications in agriculture, forestry and food sciences (for example in plant breeding), values of 5 and upward are suitable for quality control analysis, and values of above 8 are excellent, and can be used in any analytical situation.

MATERIALS AND METHODS

Materials

Ten trees from five clones of 4-year-old eucalyptus hybrids, planted in Kanchanaburi Province, Thailand, were selected for this study. Wood discs from each tree were sampled at 6 tree positions, i.e. 0%, 20%, 40%, 60%, 80%, and 100% of the total stem height. The data of the tree height in each eucalyptus hybrid is expressed in Table 2 and as tree height diagram in Figure 17.

Each wood disc was divided into top and bottom discs. The top disc was chipped into small-sized wood pieces which were further prepared for analysis of fiber length and fiber coarseness. Some of the small-sized wood pieces were ground by Thomas-Wiley mill and screened in a vibratory sieving apparatus to obtain 40-60 mesh wood meal fractions for the analysis of S/G ratio of lignin as well as NIR spectra acquisition. The bottom disc was cut into 4 pieces for wood basic density determination. The pictures of the sampled eucalyptus trees, small-sized wood pieces, and wood powder of eucalyptus samples are shown in Figure 18.

Table 2 Tree height data at the percent height level

% Height level		0%	20%	40%	60%	80%	100%
SF1	Tree1	0	3.2	6.4	9.5	12.7	15.9
	Tree2	0	3.3	6.5	9.8	13.0	16.3
SF2	Tree1	0	1.9	3.8	5.6	7.5	9.4
	Tree2	0	2.0	4.0	6.0	7.9	9.9
SF3	Tree1	0	3.5	7.1	10.6	14.1	17.6
	Tree2	0	3.3	6.7	10.0	13.3	16.7
SF4	Tree1	0	3.1	6.2	9.3	12.4	15.5
	Tree2	0	3.1	6.2	9.3	12.4	15.5
SF5	Tree1	0	3.2	6.5	9.7	13.0	16.2
	Tree2	0	3.1	6.1	9.2	12.2	15.3

Remark: unit: meter

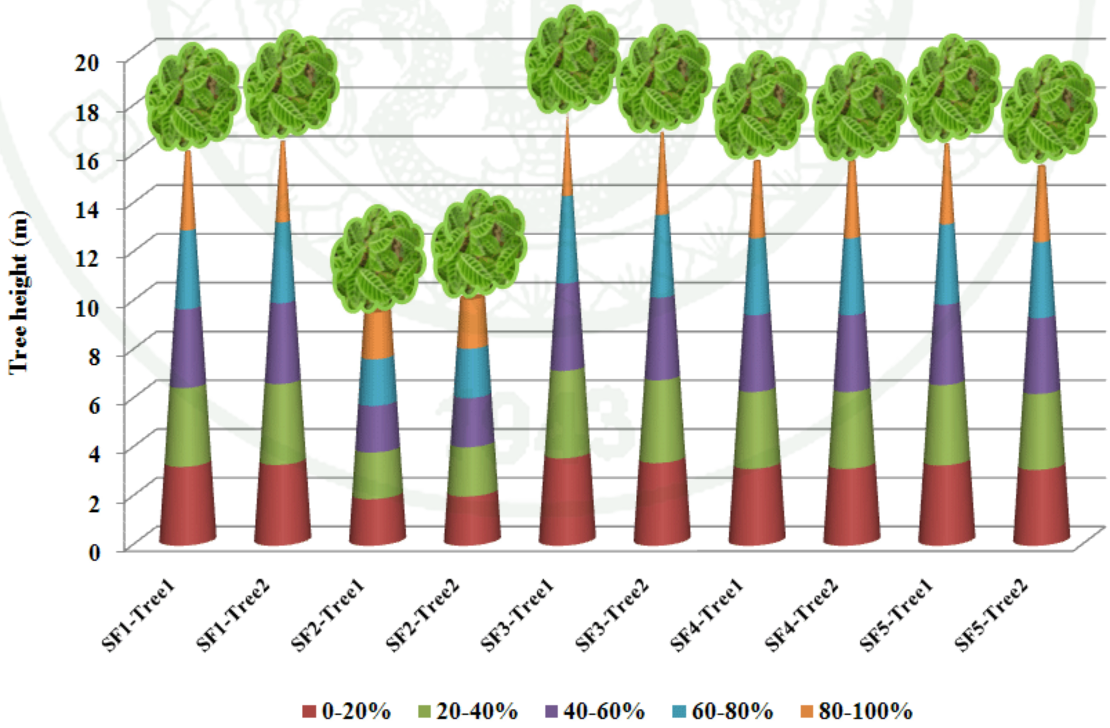


Figure 17 Tree height diagram

Figure 18 Eucalyptus trees, small-sized wood pieces, and wood powder of eucalyptus samples

Methods

1. NIR Measurements

The NIR spectra of the wood samples were acquired on a Büchi NIRFlex solids fourier transform near infrared (FT-NIR) spectrometer. A standard closed-cup was used as a sample holder for diffuse reflectance scan mode. Commonly application of 32 scans at spectral resolution of 4 cm^{-1} over the wavenumber 10000-4000 cm^{-1} (or wavelength 1000-2500 nm) was used. Conditioning room for NIR measuring was set at 25˚C and 50 % relative humidity. Figure 19 shows NIR measuring at the conditioning room, sample cell, and FT-NIR spectrophotometer.

Figure 19 Measurement of wood NIR spectra using diffuse reflectance FT-NIR spectrophotometer

2. Determination of Wood Basic Density

Wood samples were determined the basic density according to TAPPI standard procedure (T 258-om02: basic density and moisture content of pulpwood). The eucalyptus wood discs for determination of wood basic density are shown in Figure 20.

Figure 20 Eucalyptus wood discs for determination of wood basic density

3. Determination of Syringyl/Guaiacyl (S/G) Ratio of Eucalyptus Wood Lignin

The S/G ratio of eucalyptus wood lignin was analyzed by analytical pyrolysis-gas chromatography/mass spectrometry (Py-GC/MS) technique. About 80 µg of fine wood powder was pyrolyzed at 500˚C for 10 s. The pyrolysis products were directly transferred to GC Agilent 6890 for separation and MS Agilent 5973N MSD was used for identification of the S- and G- lignin units. The S/G values were calculated by using total ion chromatogram (TIC) from MS signal. The detail of S/G ratio analysis is described in Appendix A.

4. Determination of Fiber Length and Fiber Coarseness

The small-sized wood pieces were mild macerated in a 1:2 (v/v) mixture of glacial acetic acid and 6% hydrogen peroxide (according to Franklin’s method) at 70˚C for approx. 48-52 hours, and then washed with distillated water, fiberized and dried in room temperature. The length and coarseness of fibers were measured by Kajaani FS-300 Analyzer. The schematic of fiber analysis is shown in Figure 21.

Figure 21 Maceration of the eucalyptus wood fiber and measuring fiber length and fiber coarseness using Kajaani FS300 Fiber Analyzer

5. Data Processing: Multivariate Data Analysis

Unscrambler V.9.7 CAMO software was used for data pre-processing and for the calibration and validation of the PCR and PLSR models. NIR spectra were pre-processed by many mathematic methods such as 3 points smoothing filter and a second-order polynomial to obtain second derivatives (2^{nd} Der.), multiplicative scatter correction (MSC), maximum normalization (Max. Norm.), linear baseline correction (LBC), and standard normal variate (SNV) and so on (see Appendix B). The NIR evaluation were applied in whole-NIR-wavelength (1000-2500 nm) and long-NIR-wavelength (1400-2500 nm) as well for the calculation of the PCR and PLSR models.

For calibration and validation, the NIR data sets were regressed against the property of interest. Calibration and validation were evaluated by the coefficient of determination (R^2 for calibration set and adjusted R^2 for validation set) between predicted and measured values, and root mean square error of calibration/prediction (RMSEC/RMSEP), standard error of calibration/prediction (SEC/SEP), Bias, and ratio of performance to deviation (RPD_C/RPD_P), respectively. Set of total samples (N=60) were divided into calibration set (75% or N=45) and validation set (25% or N=15) when the calibration set was covered the range of validation set to perform test set validation. The schematic of the research methodology is showed in the Figure 22.

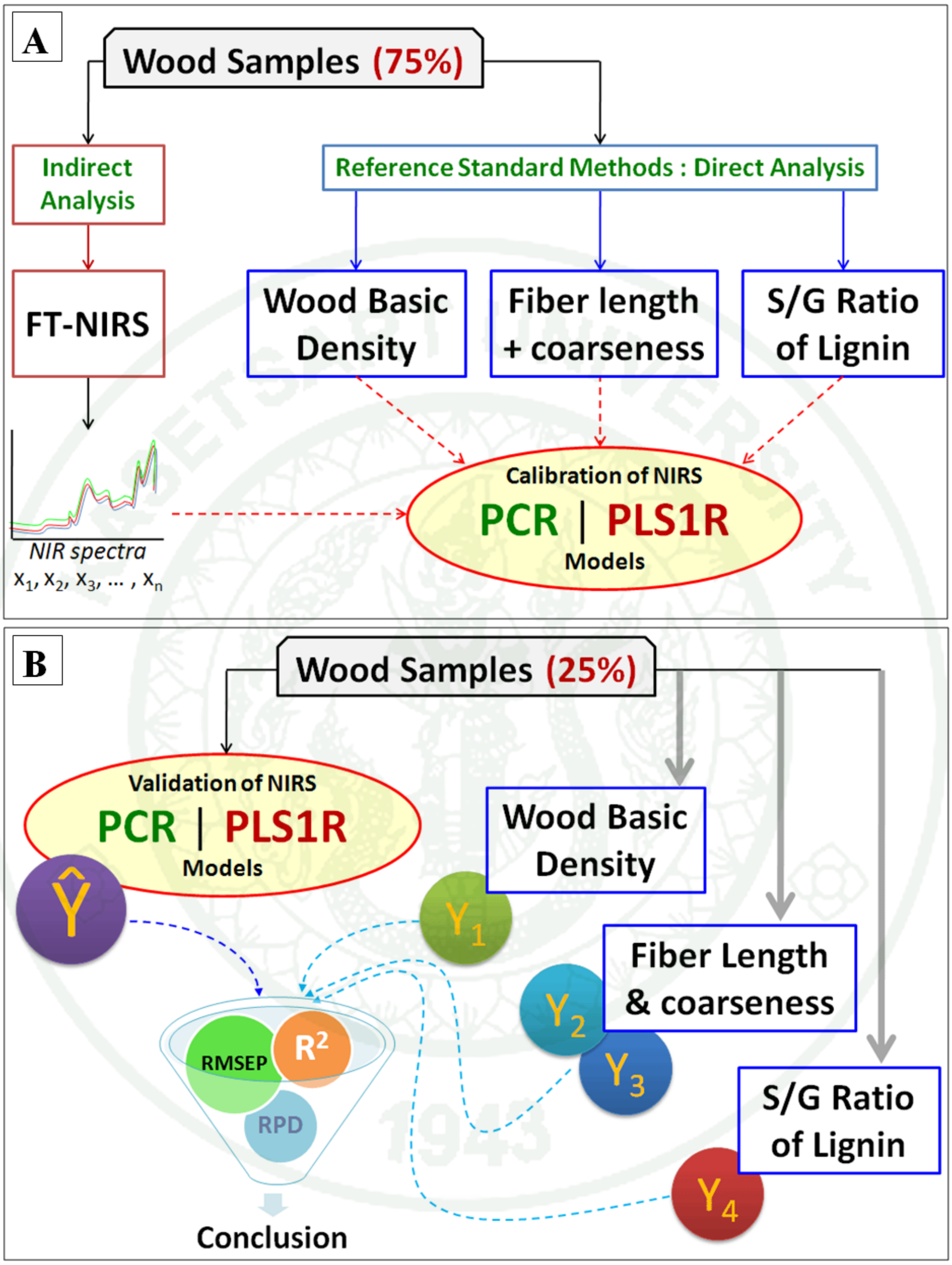


Figure 22 Schematic of the research methodology: calibration step [**A**] and validation step [**B**]

RESULTS AND DISCUSSION

1. Analysis of Wood Properties

The wood properties, i.e. wood basic density, S/G ratio of lignin, fiber length, fiber coarseness from five eucalyptus hybrids were studied in this research. The summary results, basic statistics and histogram plots of these properties are listed in Table 3, Table 4 and Figure 23, respectively.

Table 3 The summary of whole-tree wood properties of eucalyptus hybrids

Eucalyptus Hybrids	Wood basic density (kg/m^3)	S/G ratio of lignin	Fiber length (mm)	Fiber coarseness (mg/m)
SF1	659.10	2.58	0.93	0.114
SF2	646.37	2.35	0.90	0.121
SF3	589.70	2.73	0.88	0.081
SF4	527.89	2.32	0.94	0.081
SF5	629.89	2.40	0.90	0.082

Table 4 Statistical results of wood basic density, fiber length, fiber coarseness, and S/G ratio of lignin

Property	Unit	Elements	Skewness	Mean	SD
Wood basic density	kg/m^3	60	-0.44	610.59	51.99
S/G ratio of lignin	-	60	0.09	2.48	0.34
Fiber length	mm	60	-0.23	0.91	0.05
Fiber coarseness	mg/m	60	1.295	0.096	0.025

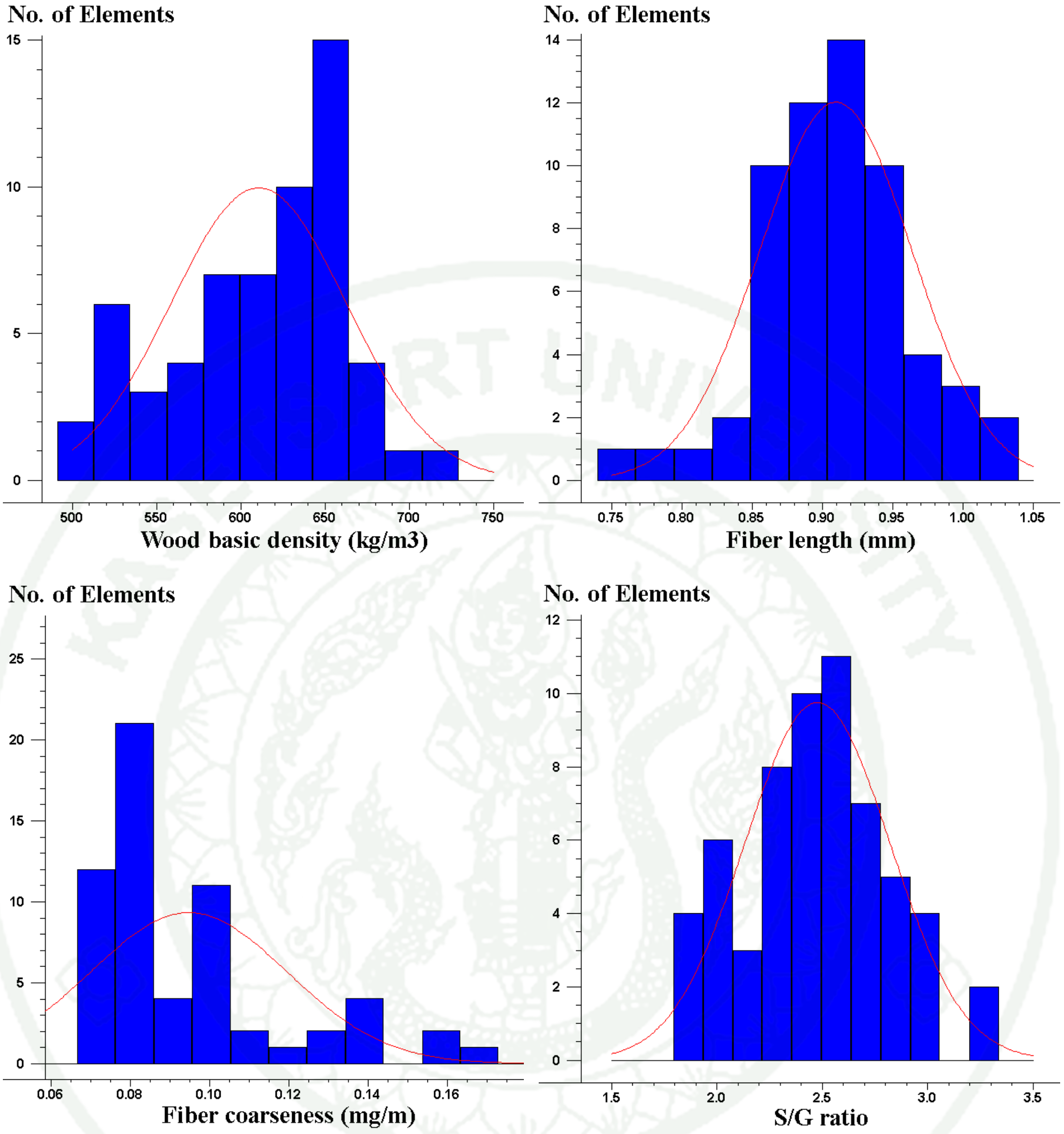


Figure 23 Histogram plots for wood basic density, S/G ratio of lignin, fiber length, and fiber coarseness

1.1 Factor Analysis of Wood Properties

For the selection of the raw material for pulp production, the factor analysis was performed from the wood properties of interest against the hybrid of eucalyptus. Table 5 shows score and loading data of factor analysis (varimax rotation) and the bi-plot, the score plot overlapped by the loading plot, is constructed as in Figure 24.

Table 5 Score and loading data of factor analysis (varimax rotation)

Score			Loading		
Samples	PC 1	PC 2	Variables	PC 1	PC 2
SF1	1.044	-0.712	Wood basic density	0.288	0.611
SF2	1.104	-0.921	S/G ratio	0.704	-0.150
SF3	0.147	2.052	Fiber length	0.554	-0.493
SF4	-2.179	-0.683	Fiber coarseness	-0.339	-0.600
SF5	-0.115	0.264			

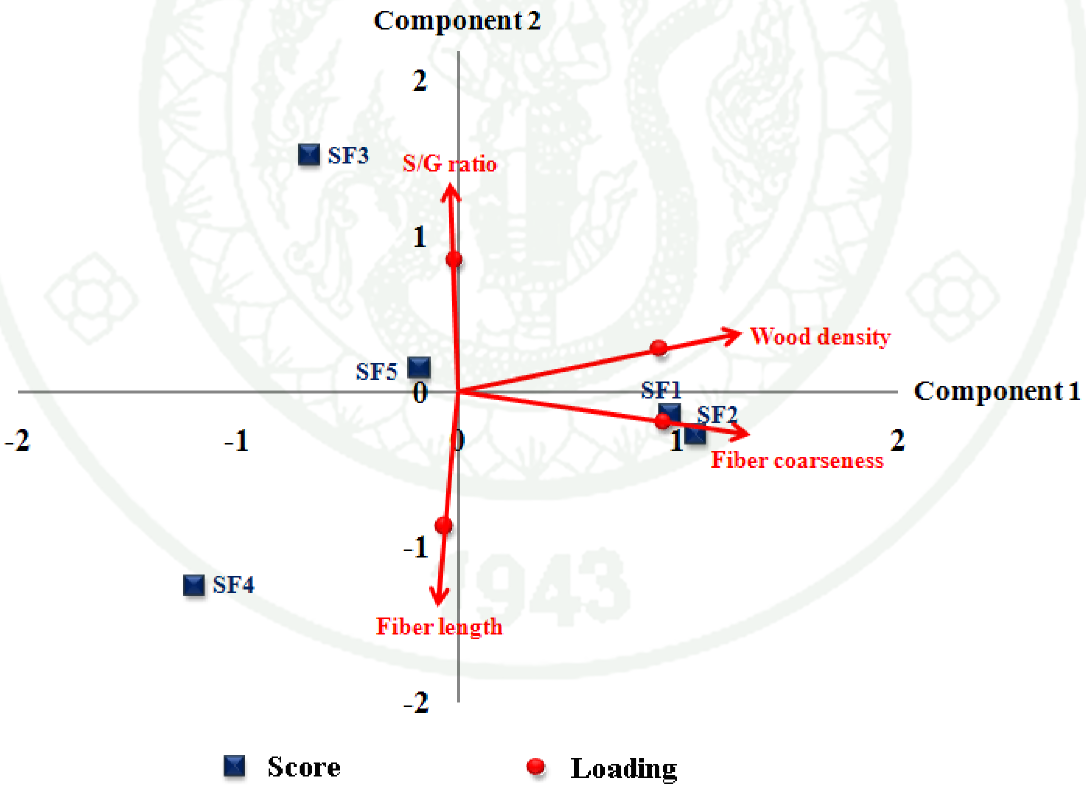


Figure 24 Bi-plot of factor analysis (varimax rotation)

From the bi-plot of factor analysis, the scores of SF1 and SF2 are response with the loading parameters of wood basic density and fiber coarseness. The score of SF3 is response with the loading parameter of lignin S/G ratio and the score of SF4 is response with the loading parameter of fiber length. The score of SF5 is moderate respond in all loading parameters.

1.2 The Correlation Analysis of Wood Properties

The correlation analysis was applied to study the relationships of wood basic density, S/G ratio of lignin, fiber length and fiber coarseness from the eucalyptus hybrids. The results from Table 6 and Figure 25 indicated the correlation between fiber coarseness and wood basic density (r=0.546, α=0.01). Its R^2 was 0.332 which means 33.2% of variance in fiber coarseness could be explained by wood basic density. The other pairs of the correlation are too low and insignificant.

Table 6 Correlation results of all investigated parameters

	Wood density	Fiber length	Fiber coarseness	S/G ratio
Wood density				
r-value		-0.152	0.546**	-0.068
Sig. (2-tailed)		0.246	0.000	0.605
N		60	60	60
Fiber length				
r-value			0.113	-0.253
Sig. (2-tailed)			0.388	0.051
N			60	60
Fiber coarseness				
r-value				0.203
Sig. (2-tailed)				0.119
N				60

Remark: Correlation **significant at the 0.01 level (2-tailed)

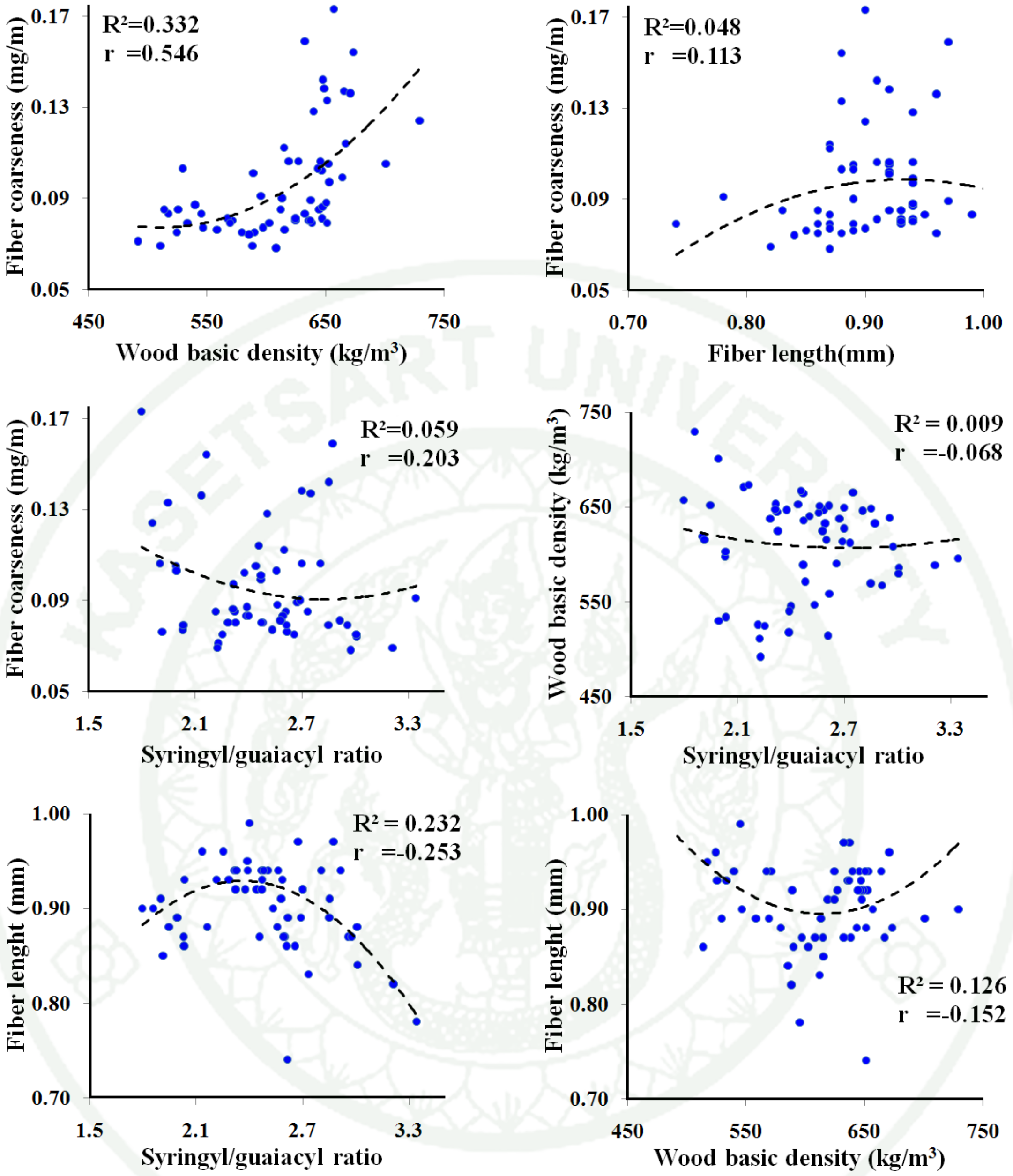


Figure 25 Correlation plots for wood basic density, S/G ratio of lignin, fiber length and fiber coarseness

2. NIR Spectral Data

2.1 Spectral measurements

The NIR spectral data were obtained in the range of 1000-2500 nm. The resolution of the adjacent wavelengths was taken as 1 nm. The total number of absorbance points was 1501 points. In NIR reflectance spectroscopy, the baseline slopes upwards with increasing wavelength due to the sample variation with particle size. Figure 26 shows the original spectra of average samples from 5 clones of 4-year-old eucalyptus hybrids wood powder in the range of 1000-2500 nm.

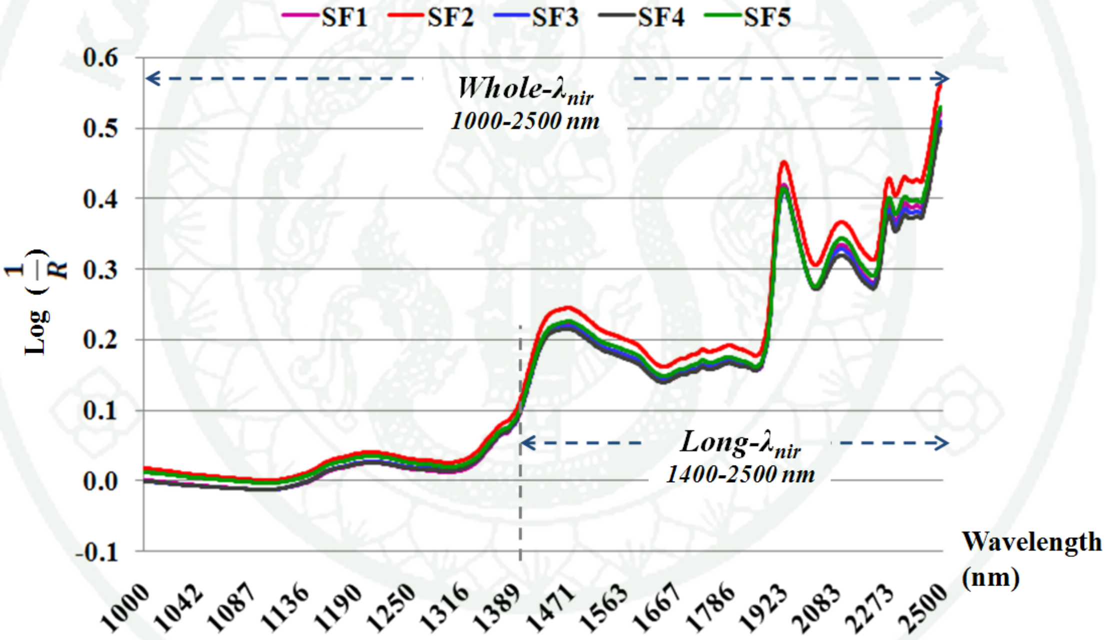


Figure 26 Near infrared spectra of eucalyptus wood

2.2 Spectral pre-processing data

In this experiment, 14 mathematic pre-processing types and combinations were used for comparison in NIR model performance (see Appendix Figure B1) as following:

- No pre-processing
- Multiplicative scatter correction
- Savitzky-Golay smoothing smoothing 3 points
- Savitzky-Golay smoothing smoothing 17 points
- Savitzky-Golay smoothing smoothing 33 points
- Vector normalization
- Maximum normalization
- Linear baseline correction + Standard normal variate
- Multiplicative scatter correction + Savitzky-Golay (smoothing 3 points)
- Multiplicative scatter correction + First derivative (smoothing 3 points)
- Multiplicative scatter correction + Second derivative (smoothing 3 points)
- First derivative (smoothing 3 points) + Vector normalization
- Second derivative (smoothing 3 points) + Vector normalization
- First derivative (smoothing 17 points) + Vector normalization

In addition to comparing NIR modeling performance in different mathematic pre-processing methods, the comparison between the use of whole-NIR-wavelength (1000-2500 nm) and only long-NIR-wavelength (1400-2500 nm) was examined in this research as well (Figure 26).

3. Multivariate Data Analysis

Spectral data were analyzed by using principal component regression (PCR) and partial least squares regression (PLSR) analysis method to develop a relationship between the value from conventional method and the value from NIR spectral data.

Data sets comprising of the NIR spectra were split into two approximate subsets used for modeling data sets; a calibration set and validation set for establishing a predictive relationship between the spectra and wood property of interest. The splitting was 45 samples for calibration set and then 15 samples for validation set. The results of basic statistics, i.e. number of sample, range of data, mean, and standard deviation of the wood property of interest separated by set of calibration and validation are shown in Table 7.

The evaluation of NIR models results varying with spectral pre-processing methods, whole-NIR and long-NIR wavelength, PCR and PLSR are shown in Appendix C: Appendix Table C1-C16.

Only the one best pre-processing of NIR spectra in each wood property of interest is selected to compare and discuss in the section 3.1 to 3.5.

Table 7 Basic statistics of wood basic density, S/G ratio of lignin, fiber length and fiber coarseness separated by set of calibration and validation

Property	Set	N	Range	Mean	SD
Wood basic density (kg/m^3)	Total	60	491.35-729.34	610.59	51.99
	Calibration	45	491.35-729.34	610.12	52.22
	Validation	15	513.86-700.76	612.01	53.09
Fiber length (mm)	Total	60	0.74-1.04	0.91	0.05
	Calibration	45	0.74-1.04	0.91	0.06
	Validation	15	0.86-1.01	0.91	0.04
Fiber coarseness (mg/m)	Total	60	0.068-0.173	0.096	0.025
	Calibration	45	0.068-0.173	0.094	0.026
	Validation	15	0.077-0.138	0.101	0.021
Syringyl/guaiacyl ratio	Total	60	1.80-3.34	2.48	0.34
	Calibration	45	1.80-3.34	2.47	0.34
	Validation	15	1.90-3.21	2.49	0.36

3.1 Analysis of spectral data and wood basic density

Multivariate regression models, PCR and PLSR, were developed for determination of wood basic density using NIR spectra as dependent variable. Wavelength range, between whole-NIR (1000-2500 nm) and long-NIR (1400-2500 nm), and pre-processing methods of NIR spectra as well as regression types were determined from several combinations to give highest coefficients of determination, lowest error for both calibration and prediction, and highest RPD (see Appendix Table C1-C4). From the results, it can be concluded that the use of whole-NIR range gave the better result for the NIR model performance to prediction than the use of only selected wavelength in long-NIR region. Moreover, the statistical result obtained from PLSR was shown better than obtained from PCR as indicated in Table 8, which expresses the statistical results separately between calibration and validation sets. The relationship between predicted and measured values for calibration and validation of wood basic density is plotted in Figure 27.

Table 8 Calibration and test-set validation results of PCR and PLSR for wood basic density

Spectral pre-processing		Non pre-processing			
NIR range		Whole-λ_{nir} 1000-2500 nm		Long-λ_{nir} 1400-2500 nm	
Regression type		PCR	PLSR	PCR	PLSR
Set	No. of factors	13	11	14	9
Calibration (N=45)	R^2	0.894	0.922	0.890	0.897
	RMSEC	16.773	14.453	17.109	16.596
	SEC	16.963	14.616	17.302	16.784
	Bias	6.78×10^{-7}	1.56×10^{-5}	8.82×10^{-6}	1.63×10^{-5}
	RPD_c	3.08	3.57	3.02	3.11
Validation (N=15)	adjusted R^2	0.881	0.901	0.869	0.857
	RMSEP	17.705	16.124	18.605	19.419
	SEP	18.321	16.659	19.250	20.095
	Bias	4.27×10^{-1}	9.92×10^{-1}	-5.20×10^{-1}	-4.31×10^{-1}
	RPD_p	2.85	3.13	2.71	2.60

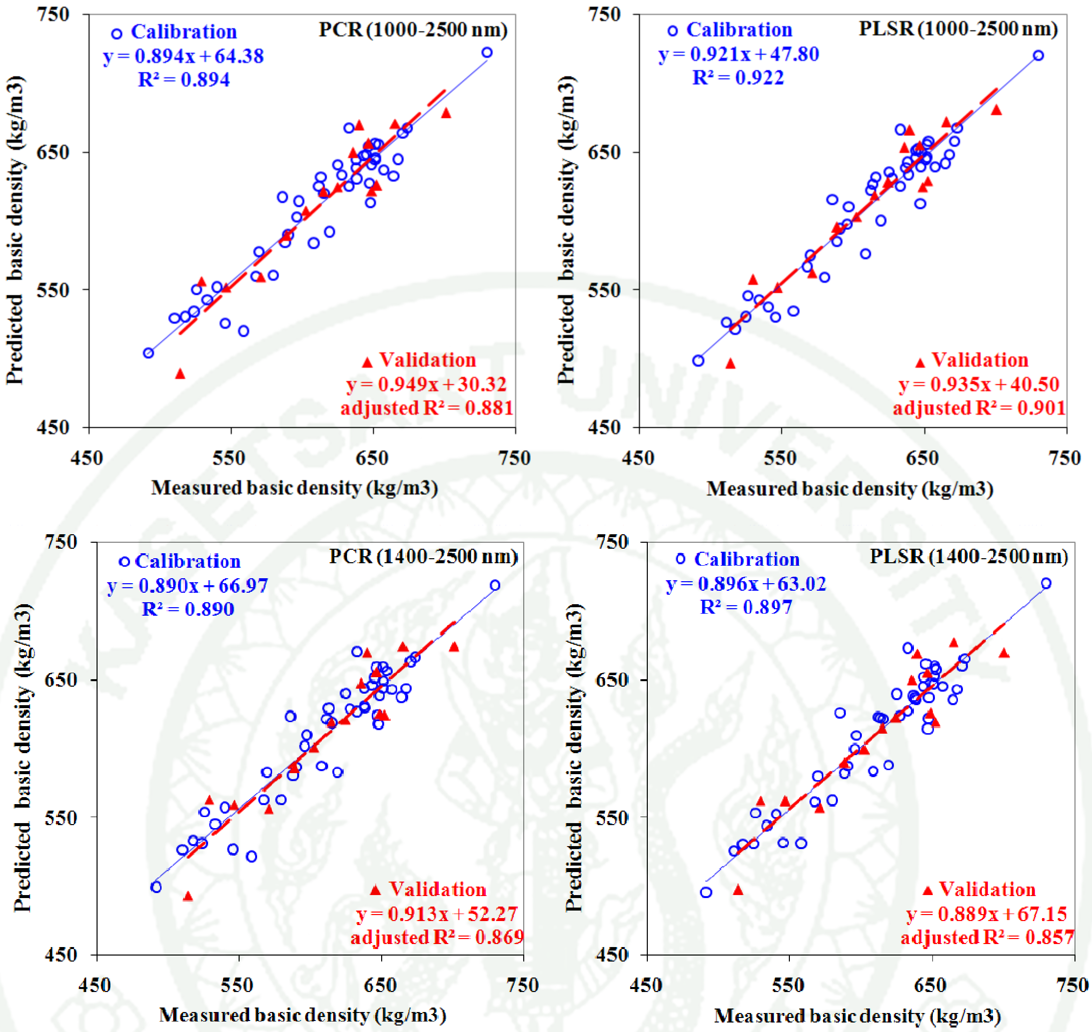


Figure 27 Predicted vs. measured value plots of wood basic density

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 highly important (significant) and low value (close to 0) indicate an unimportant variable. The regression coefficients of the non pre-processing NIR spectra between 1000-2500 nm of PCR and PLSR models for wood basic density are plotted in Figure 28. The strong regression coefficients were observed in PLSR over the PCR for the case of wood basic density that means the regression using PLS algorithm was preferred for wood basic density.

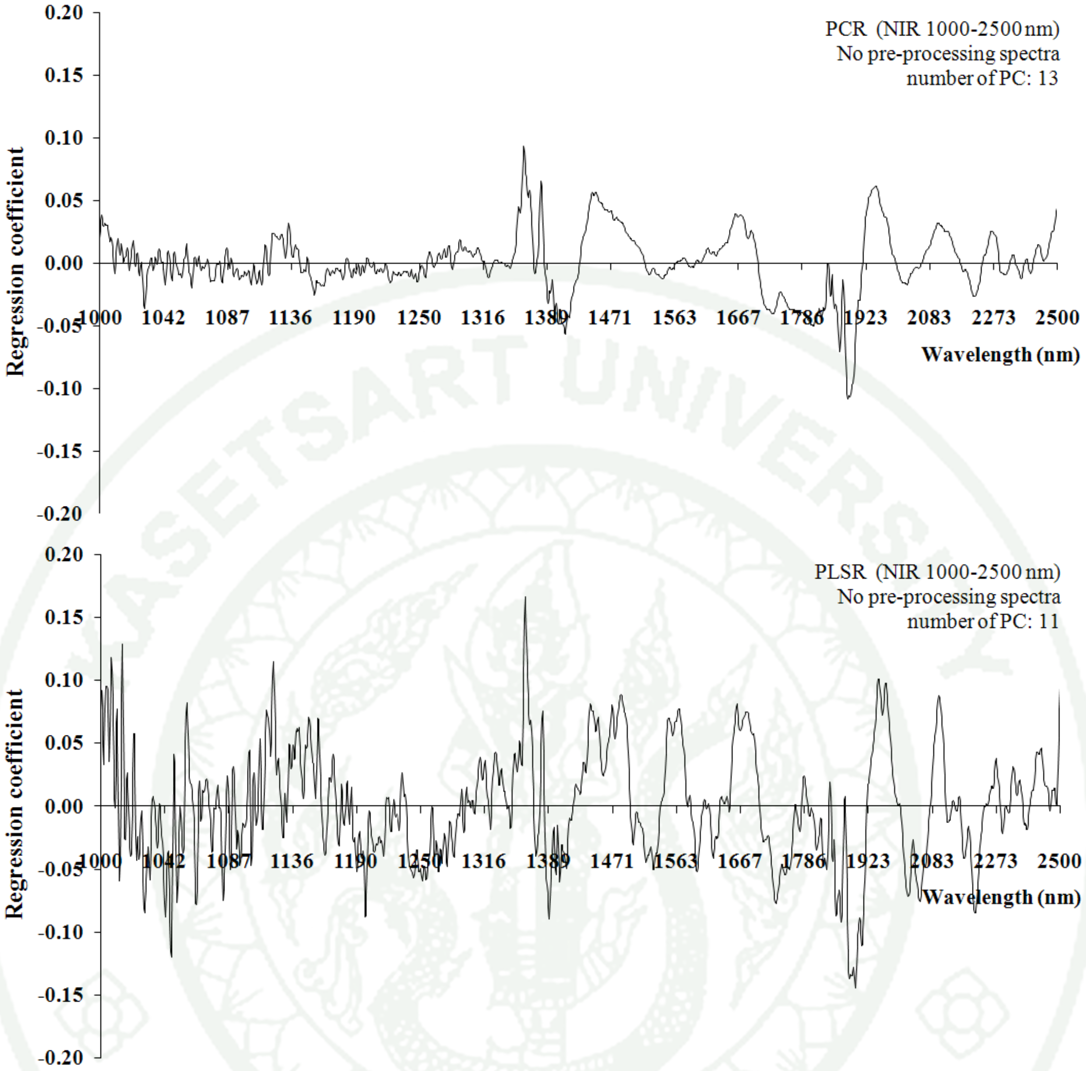


Figure 28 Regression coefficient plots of PCR and PLSR of NIR evaluation between 1000-2500 nm for wood basic density

3.2 Analysis of spectral data and S/G ratio of lignin

Multivariate regression models, PCR and PLSR, were developed for determination of S/G ratio of lignin using NIR spectra as dependent variable. Wavelength range, between whole-NIR and long-NIR, and pre-processing methods of NIR spectra as well as regression types were determined from several combinations to give highest R^2, lowest SEC and SEP, and highest RPD (see Appendix Table C5-C8). From the result, it can be concluded that the use of whole-NIR range gave the better result for the NIR model performance to prediction than the use of only selected wavelength as long-NIR region, and the use of linear baseline correction combined with standard normal variate as NIR spectral pre-processing gave the best result in prediction. The calibration and test-set validation results of PCR and PLSR for S/G ratio is shown in Table 9.

Table 9 Calibration and test-set validation results of PCR and PLSR for S/G ratio

Spectral pre-processing		LBC + SNV			
NIR range		Whole-λ_{nir} 1000-2500 nm		Long-λ_{nir} 1400-2500 nm	
Regression type		PCR	PLSR	PCR	PLSR
Set	No. of factors	12	11	11	10
Calibration (N=45)	R^2	0.942	0.974	0.941	0.969
	RMSEC	0.082	0.055	0.082	0.060
	SEC	0.083	0.055	0.083	0.060
	Bias	-1.03×10^{-7}	-1.40×10^{-7}	-1.09×10^{-7}	-1.33×10^{-7}
	RPD_c	4.15	6.23	4.13	5.71
Validation (N=15)	adjusted R^2	0.858	0.870	0.849	0.864
	RMSEP	0.130	0.125	0.134	0.128
	SEP	0.134	0.129	0.139	0.132
	Bias	8.71×10^{-3}	3.68×10^{-3}	1.83×10^{-3}	2.71×10^{-3}
	RPD_p	2.57	2.67	2.48	2.61

According to Table 9, the statistical result obtained from PLSR was shown better than obtained from PCR. The relationship between predicted and measured values for calibration and validation of S/G ratio of lignin is plotted in Figure 29.

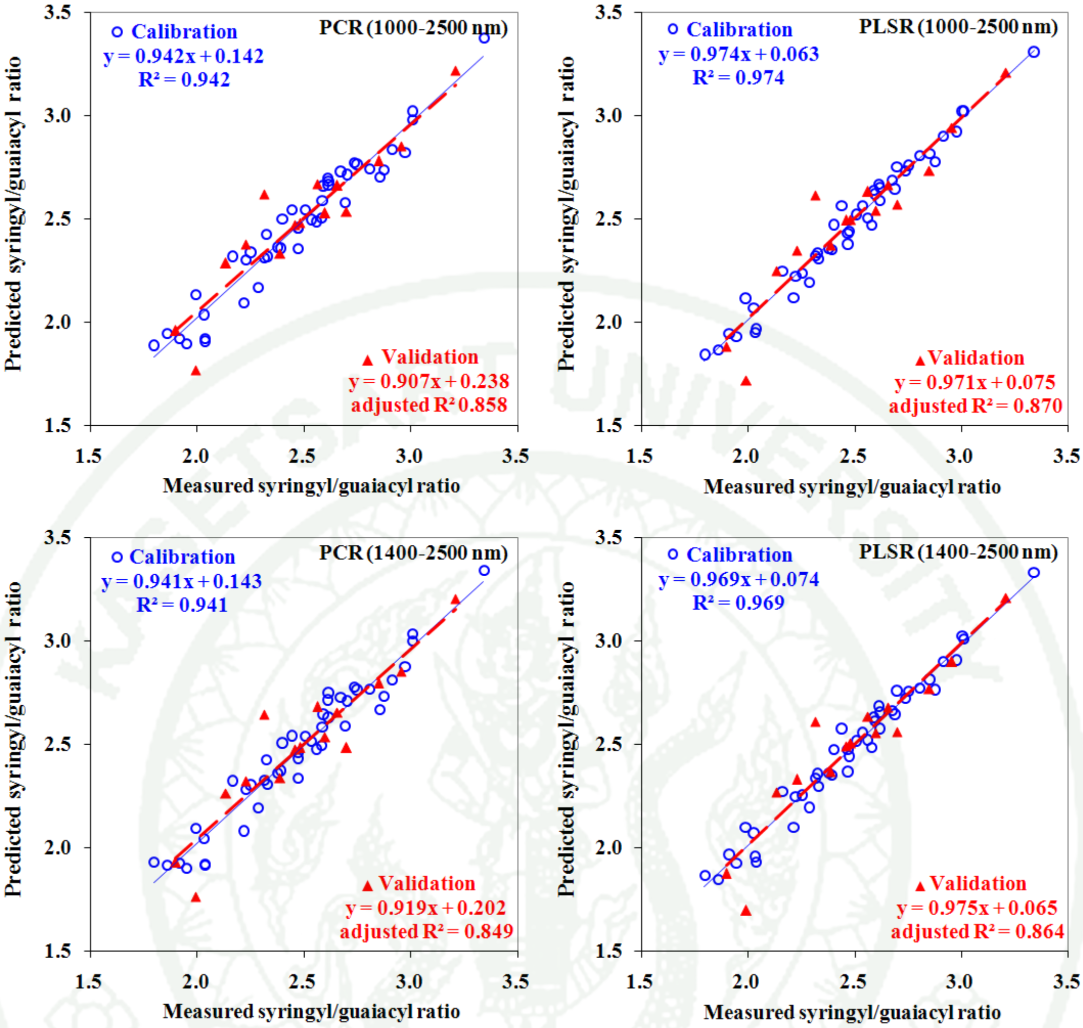


Figure 29 Predicted vs. measured value plots of syringyl/guaiacyl ratio

The regression coefficients of linear baseline correction combined with standard normal variate pre-processing NIR spectra between 1000-2500 nm of PCR and PLSR models for S/G ratio of lignin are plotted in Figure 30. The strong regression coefficients were observed in PLSR over the PCR for the case of S/G ratio that means the regression using PLS algorithm was preferred for S/G ratio.

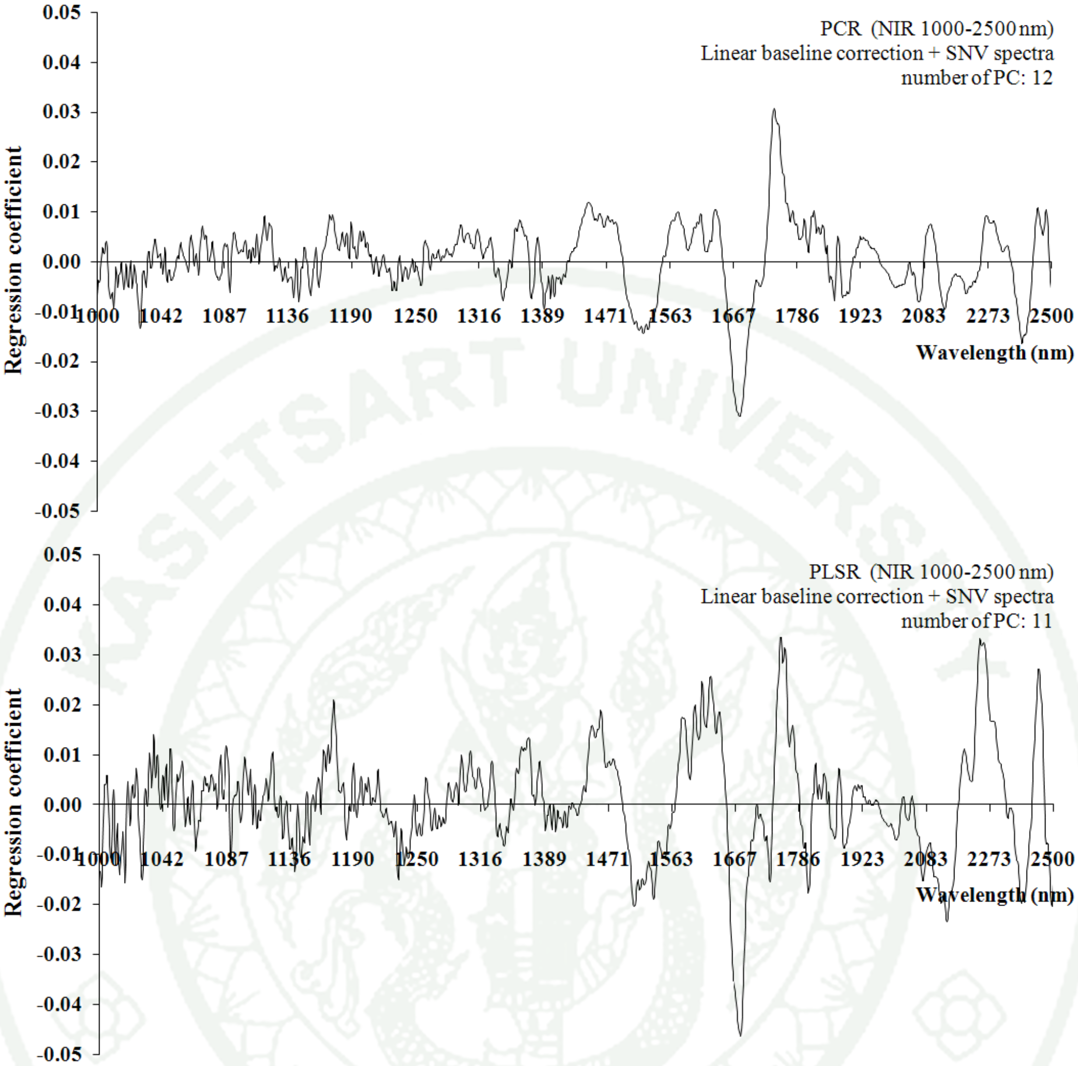


Figure 30 Regression coefficient plots of PCR and PLSR of NIR evaluation between 1000-2500 nm for syringyl/guaiacyl ratio of lignin

3.3 Analysis of spectral data and fiber length

Multivariate regression models, PCR and PLSR, were developed for determination of fiber length using NIR spectra as dependent variable. Wavelength range, between whole-NIR (1000-2500 nm) and long-NIR (1400-2500 nm), and pre-processing methods of NIR spectra as well as regression types were determined from several combinations to give highest coefficients of determination, lowest error for both calibration and prediction, and highest RPD (see Appendix Table C9-C12).

From the results, it can be concluded that the use of whole-NIR range gave the better result for the NIR model performance to prediction than the use of only selected wavelength as long-NIR region, and the use of multiplicative scatter correction combined with the second derivative 3 smoothing point (MSC + 2^{nd} Der. (3 pts)) as NIR spectral pre-processing gave the best result in prediction. Moreover, the statistical result obtained from PLSR was shown better than obtained from PCR as indicated in Table 10, which expresses the statistical results separately between calibration and validation sets. The relationship between predicted and measured values for calibration and validation of fiber length is plotted in Figure 31.

Table 10 Calibration and test-set validation results of PCR and PLSR for fiber length

Spectral pre-processing		MSC + 2^{nd} Der. (3 pts)			
NIR range		Whole-λ_{nir} 1000-2500 nm		Long-λ_{nir} 1400-2500 nm	
Regression type		PCR	PLSR	PCR	PLSR
Set	No. of factors	8	3	5	5
Calibration (N=45)	R^2	0.473	0.836	0.376	0.952
	RMSEC	0.040	0.022	0.044	0.012
	SEC	0.040	0.023	0.044	0.012
	Bias	4.50×10^{-8}	4.77×10^{-8}	3.71×10^{-8}	4.11×10^{-8}
	RPD$_c$	1.38	2.47	1.27	4.59
Validation (N=15)	adjusted R^2	0.594	0.623	0.586	0.591
	RMSEP	0.032	0.031	0.032	0.032
	SEP	0.032	0.031	0.031	0.031
	Bias	8.43×10^{-3}	8.25×10^{-3}	1.12×10^{-2}	1.02×10^{-2}
	RPD$_p$	1.76	1.83	1.79	1.78

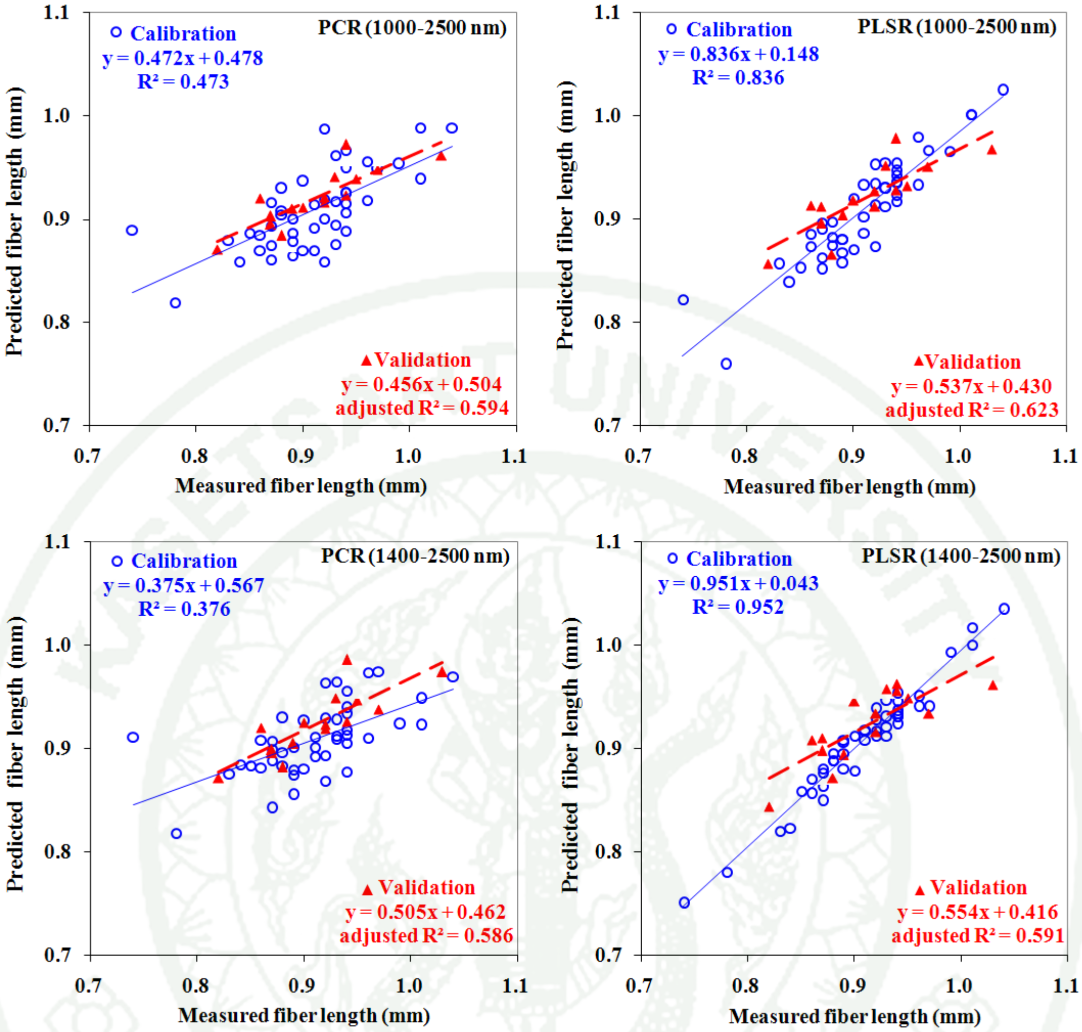


Figure 31 Predicted vs. measured value plots of fiber length

The PLSR was preferred to use for developing the NIR model for assessment of the fiber length, as indicated the great in statistical results in Table 10, however, this NIR model of fiber length was unsuitable to use because the slope of the validation line was different from that of the calibration line. Moreover, both calibration and validation lines differed from the target line (45-degree line) as well (Figure 31).

The regression coefficients of the MSC + 2nd Der. (3 pts) pre-processing NIR spectra between 1000-2500 nm of PCR and PLSR models for fiber length are plotted in Figure 32. The strong regression coefficients were observed in PLSR over the PCR for the case of fiber length that means the regression using PLS algorithm was preferred for fiber length. However, the coefficients are look like noise.

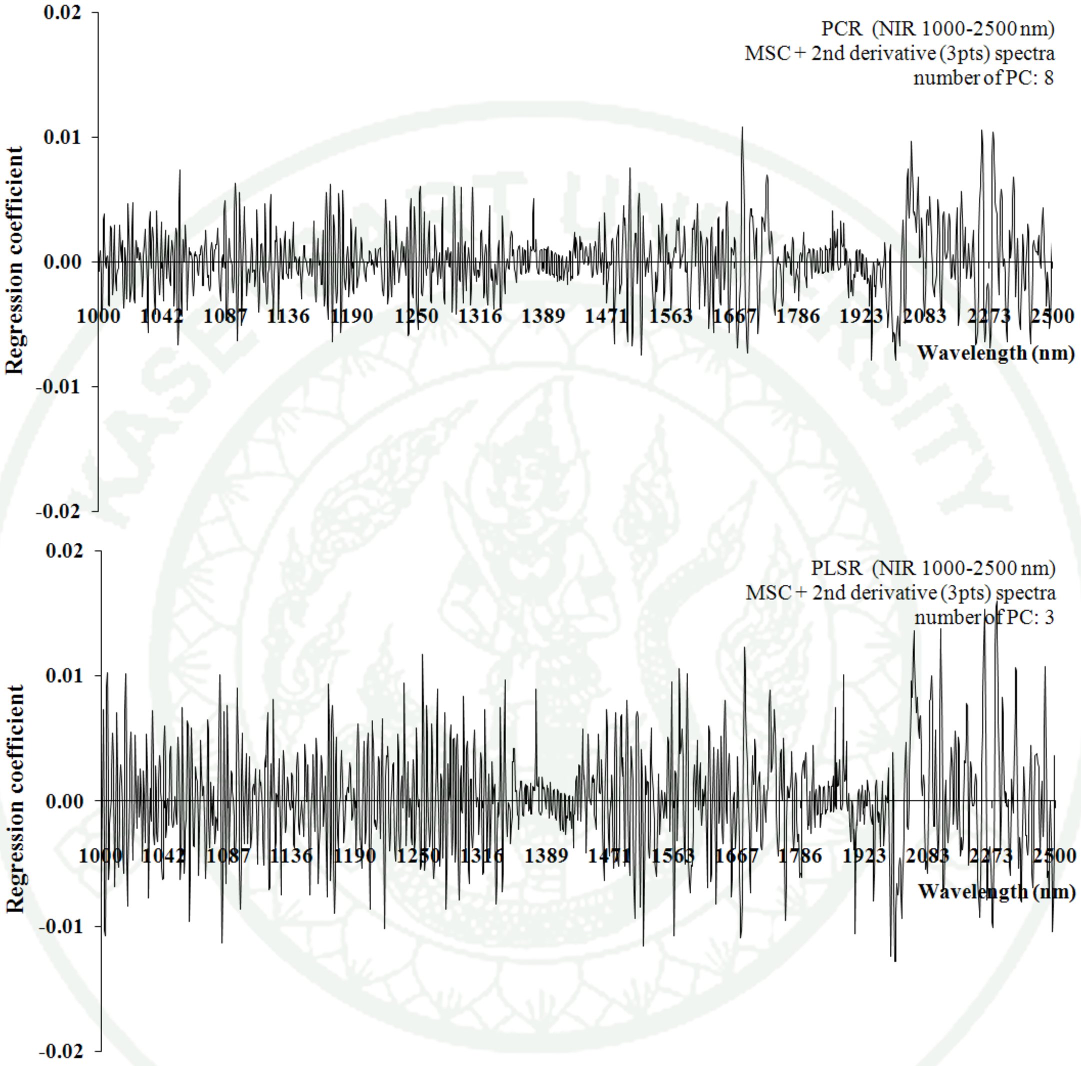


Figure 32 Regression coefficient plots of PCR and PLSR of NIR evaluation between 1000-2500 nm for fiber length

3.4 Analysis of spectral data and fiber coarseness

Multivariate regression models, PCR and PLSR, were developed for determination of fiber coarseness using NIR spectra as dependent variable. Wavelength range, between whole-NIR (1000-2500 nm) and long-NIR (1400-2500 nm), and pre-processing methods of NIR spectra as well as regression types were determined from several combinations to give highest coefficients of determination, lowest error for both calibration and prediction, and highest RPD (see Appendix Table C13-C16). From the result, it can be concluded that the use of whole-NIR range gave the better result for the NIR model performance to prediction than the use of only selected wavelength as long-NIR region, and the use of maximum normalization (Max. Norm.) as NIR spectral pre-processing gave the best result in prediction. In opposite to the case of wood basic density and fiber length, for fiber coarseness the statistical result obtained from PCR is shown better prediction performance than obtained from PLSR as indicated in Table 11, which expresses the statistical results separately between calibration and validation sets. The relationship between predicted and measured values for calibration and validation of fiber coarseness is plotted in Figure 33.

Table 11 Calibration and test-set validation results of PCR and PLSR for fiber coarseness.

Spectral pre-processing		Max. Norm.			
NIR range		Whole-λ_{nir} 1000-2500 nm		Long-λ_{nir} 1400-2500 nm	
Regression type		PCR	PLSR	PCR	PLSR
Set	No. of factors	17	8	16	9
Calibration (N=45)	R^2	0.751	0.682	0.720	0.793
	RMSEC	0.012	0.014	0.013	0.011
	SEC	0.012	0.014	0.013	0.011
	Bias	-1.48×10^{-7}	2.34×10^{-7}	-2.78×10^{-7}	-1.96×10^{-7}
	RPD_c	2.00	1.77	1.89	2.20
Validation (N=15)	adjusted R^2	0.736	0.689	0.611	0.613
	RMSEP	0.013	0.014	0.016	0.016
	SEP	0.013	0.014	0.016	0.016
	Bias	2.10×10^{-4}	8.07×10^{-4}	-5.47×10^{-4}	-1.10×10^{-3}
	RPD_p	1.97	1.82	1.62	1.63

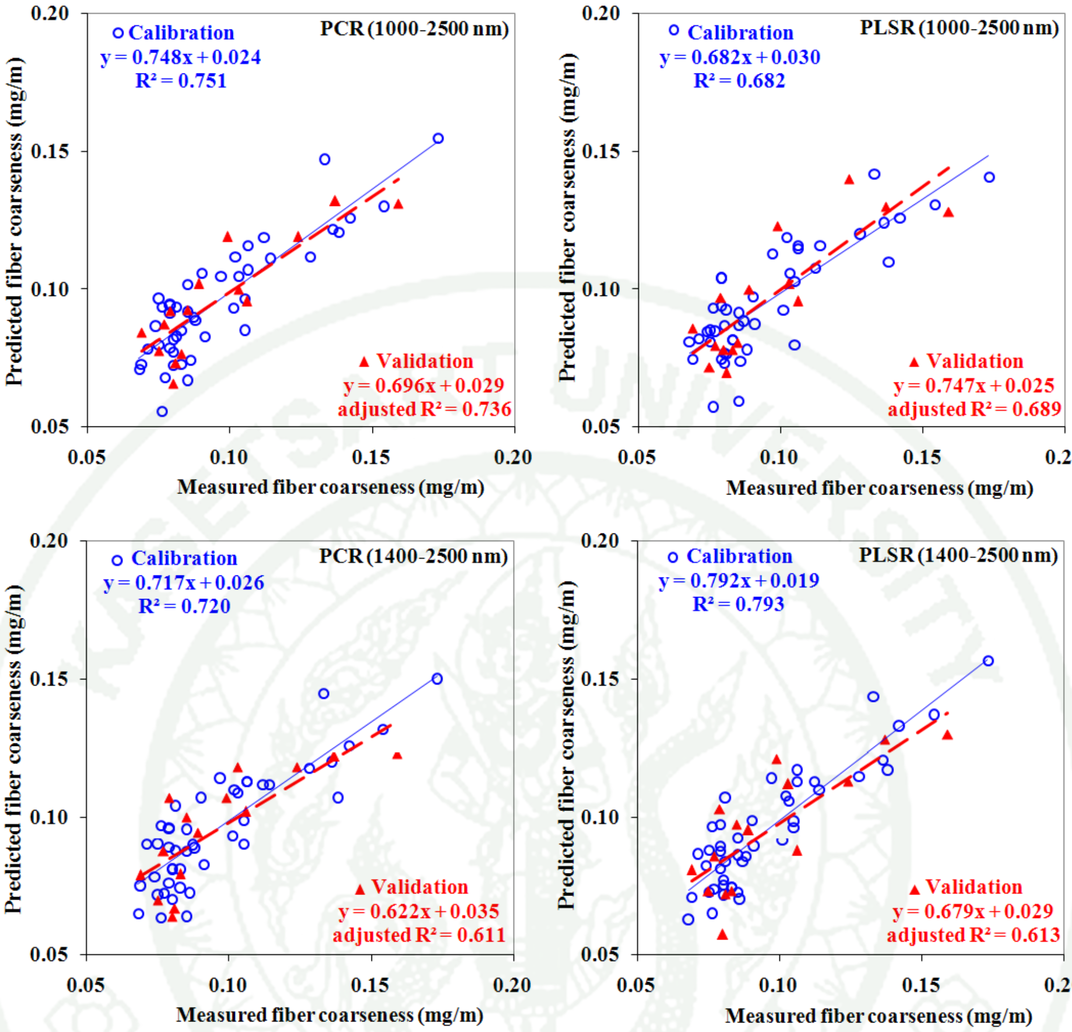


Figure 33 Predicted vs. measured value plots of fiber coarseness.

The PCR was preferred to use for developing the NIR model for assessment the fiber coarseness, as indicated the great in statistical results in Table 11, however, the application of this NIR model must be careful. From the correlation plot between predicted vs. measured values in Figure 33 and the histogram plot of fiber coarseness data which was not normal distribution (Figure 23) gave a point to mention that this model needed to be improved.

The regression coefficients of the maximum normalization pre-processing NIR spectra between 1000-2500 nm of PCR and PLSR models for fiber coarseness are plotted in Figure 34. The strong regression coefficients were observed in PCR over the PLSR especially in the long-NIR region for the case of fiber coarseness.

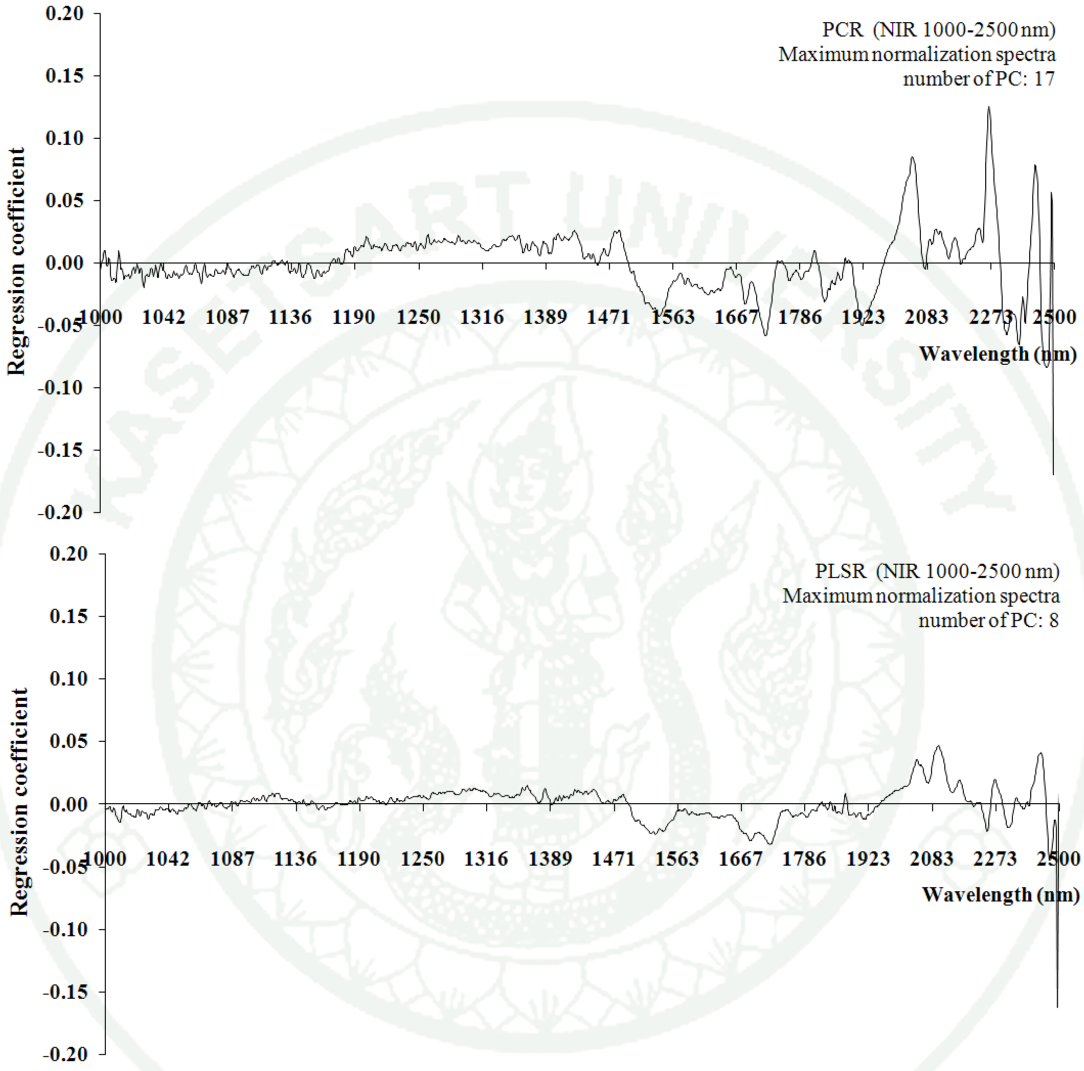


Figure 34 Regression coefficient plots of PCR and PLSR of NIR evaluation between 1000-2500 nm for fiber coarseness.

3.5 NIR analysis

The results in Table 8 to 11 show that calibration performances are generally better than validation performances due to the principals of multivariate regression analysis, especially PLSR. The use of PLS to regress wood basic density, fiber length and S/G ratio of lignin were effective, except for fiber coarseness which gave a good result when it was regressed by PCR.

For the validation of test set, high adjusted R^2 between predicted and measured values were obtained from wood basic density (adjusted R^2=0.901) and S/G ratio of lignin (adjusted R^2=0.870), and moderate-high adjusted R^2 were obtained from fiber coarseness and fiber length (adjusted R^2=0.736 and 0.623, respectively). Length and coarseness of eucalyptus fibers were difficult to predict which may be caused by the range of fiber length and fiber coarseness of eucalyptus wood were much smaller than that of softwood tracheids.

In several studies, non-destructive estimation by NIR spectroscopy for eucalyptus density was evaluated. The R^2 as well as RMSEP obtained from this research were comparable with the results of Inagaki *et al.* (2010) and Ruiz *et al.* (2005), and were higher than those obtained from Ricardo and Hein (2010). However, these statistical evaluations from different researches could not exactly compare among each other, due to differences of number/set of samples, instrument, standard measurement and so on.

As the standard deviations of wood parameters were different in each study, the RPD values were calculated to take this into account and to be able to compare the PCR and PLS regression results with those from previous studies. From the analytical point of view, the RPD should be in the following range: RPD ≥ 2.5 for screening in breeding programs; RPD ≥ 5.0 acceptable for quality control; RPD ≥ 8.0 good for process control, development and applied research (Alves *et al.*, 2010).

The results of this study showed that RPD_P = 3.13 for wood basic density and RPD_P = 2.67 for S/G ratio of lignin allowed drawing the conclusion that these models were at least applicable for screening in tree breeding programs as they lied in-between screening (RPD ≥ 2.5) and quality control (RPD ≥ 5.0). The RPD values for prediction of eucalyptus density obtained in previous studies ranged from 1.6 (Ricardo and Hein, 2010) up to 3.8 (Inagaki *et al.*, 2010). In the case of fiber characteristics, RPD_P of fiber length and fiber coarseness (1.83 and 1.97, respectively) in this study were not high enough to use for screening in tree breeding programs. Therefore, improvement of the model statistics and selection of NIR wavelength, and/or increasing number of samples for training and testing the models should be further investigated. Inagaki *et al.* (2010) revealed that the RPD values for prediction of eucalyptus fiber length ranged in between 2.8-3.9. The result from Jones *et al.* (2005) showed that the RPD values for prediction of loblolly pine (*Pinus taeda* L.) tracheid coarseness ranged between 0.81-2.10, respectively.

CONCLUSION AND RECOMMENDATION

Conclusion

Average of wood basic density, lignin S/G ratio, fiber length and fiber coarseness of ten trees from five clones of 4-year-old eucalyptus hybrids were 610.6 kg/m^3, 2.48, 0.91 mm and 0.957 mg/m, respectively. The results from correlation analysis of wood properties indicated the correlation between fiber coarseness and wood basic density.

It can be concluded from NIR analysis that the use of whole-NIR range gave the better NIR model performance than the use of only selected wavelength in long-NIR region. The best NIR models of wood basic density, S/G ratio and fiber length were obtained from PLSR, whereas that of fiber coarseness was obtained from PCR.

The results of this study demonstrated that NIR spectroscopy has been successful to use as a non-destructive tool to estimate the wood basic density and S/G ratio of eucalyptus hybrids with accuracy and precision enough for screening in tree breeding programs (RPD ≥ 2.5). However, the use of NIR technique for fiber characteristics of eucalyptus hybrids is still need to be improved.

Recommendation

Study more about the relation between phenomenons of NIR spectra and characteristics of wood fiber is the challenge. Indirect relationship between fiber length and coarseness to the NIR absorbance is still need to clarify.

Multivariate data analysis (chemometrics) is very important for study the feasibility to use of NIR technology in wood and paper science. Selective range of wavelength before regression method such as moving window partial least square regression (MWPLSR) is also interesting. Advance study in chemometrics will be getting the result better in accuracy and precision due to more understanding to extract the valuable data from NIR spectra for modeling.

LITERATURE CITED

Alves, A., M. Schwanninger, H. Pereira and J. Rodrigues. 2006a. Calibration of NIR to assess lignin composition (H/G ratio) in maritime pine wood using analytical pyrolysis as the reference method. **Holzforschung.** 60: 29-31.

______, ______, ______ and ______. 2006b. Analytical pyrolysis as a direct method to determine the lignin content in wood – Part 1: Comparison of pyrolysis lignin with Klason lignin. **J. Anal. Appl. Pyrolysis.** 76: 209-213.

______, J. Rodrigues, R. Wimmer and M. Schwanninger. 2008. Analytical pyrolysis as a direct method to determine the lignin content in wood – Part 2: Evauation of the common model and the influence of compression wood. **J. Anal. Appl. Pyrolysis.** 81: 167-172.

______, A. Santos, P. Rozenburg, L.E. Pâques, J.-P. Charpentier, M. Schwanninger, and J. Rodrigues. 2010. A common near infrared-based partial least squares regression model for the prediction of wood density of *Pinus pinaster* and *Larix* × *eurolepis*. **Wood Sci. Technol.** Published online: 05 October 2010. *Cited* AACC. 1999. **Near-infrared methods-guidelines for model development and maintenance**. American Association of Cereal Chemists (AACC). AACC Method 39-00:15.

Baillères, H, F. Davrieux and F. Ham-Pichavant. 2002. Near infrared analysis as a tool for rapid screening of some major wood characteristics in a eucalyptus breeding program. **Ann. For. Sci.** 59: 479-490.

Biermann, C.J. 1996. **Handbook of Pulping and Papermaking.** Academic Press, San Diego.

Bose, S.K., R.C. Francis, M. Govender, T. Bush and A. Spark. 2009. Lignin content versus syringyl to guaiacyl ratio amongst poplars. **Bioresource Technology.** 100: 1628-1633.

Bremer, J. 1991. **Quantifizierung der gerüstsubstanzen von lignocellulosen durch analytische pyrolyse-gasschromatographie/massenspectrometrie**. Ph.D. Thesis. University of Hamburg, Fachbereich Biologie.

Brun, F.L., M. Hinchee, F.S. Gomes and N. Ramsey. n.d. **Potentiality of genetically modified trees for the eucalypt kraft pulp industry.** Available Source: http://www.celuloseonline.com.br/imagembank/Docs/DocBank/Eventos/430/4BrunOral.pdf, May 14, 2009.

Chen, C., M. Baucher, J.H. Christensen and W. Boerjan. 2001. Biotechnology in trees: Towards improved paper pulping by lignin engineering. **Euphytica.** 118: 185-195.

Collins, D. J., C. A. Pilotti and A. F. A. Wallis. 1990. Correlation of chemical composition and kraft pulping properties of some Papua New Guinea reforestation woods. **Appita J.** 43 (3): 193-198.

Davidson, J. 1996. **Ecological aspects of eucalyptus plantations.** Available Source: http://www.fao.org/docrep/005/ac777e/ac777e00.htm, May 20, 2011.

Davies, T. 1998. The history of near infrared spectroscopic analysis: Past, present and future – “From sleeping technique to the morning star of spectroscopy”. **Analusis Mag.** 26: 17-19.

Del Río J.C., A.T. Martínez and A. Gutiérrez. 2007a. Presence of 5-hydroxyguaiacyl units as native lignin constituents in plants as seen by Py-GC/MS. **J. Anal. Appl. Pyrolysis.** 79: 33-38.

Del Río J.C., A. Gutiérrez, I.M. Rodríguez, D. Ibarra and A.T. Martínez. 2007b. Composition of non-woody plant lignins and cinnamic acids by Py-GC/MS, Py/TMAH and FT-IR. **J. Anal. Appl. Pyrolysis.** 79: 39-46.

______, A. Gutiérrez, M. Hernando, P. Landín, J. Romero and A.T. Martínez. 2005. Determining the influence of eucalypt lignin composition in paper pulp yield using Py-GC/MS. **J. Anal. Appl. Pyrolysis.** 74: 110-115.

Gonzáles-Vila, F.J., G. Almendros, J.C. del Río, F. Martin, A. Gutíérrez and J. Romero. 1999. Ease of delignification assessment of wood from different Eucalyptus species by pyrolysis (TMAH)-GC/MS and CP/MAS 13C-NMR spectrometry. **J. Anal. Appl. Pyrolysis.** 49: 295-305.

Greenwood, P.F., J.D.H. van Heemst, E.A. Guthrie and P.G. Hatcher. 2002. Laser micropyrolysis GC/MS of lignin. **J. Anal. Appl. Pyrolysis.** 62: 365-373.

Guerra, A., J.P. Elissetche, M. Norambuena, J. Freer, S. Valenzuela, J. Rodríguez and C. Balocchi. 2008. Influence of lignin structural features on *Eucalyptus globulus* kraft pulping. **Ind. Eng. Chem. Res.** 47: 8542-8549.

Hodge G.R. and W.C. Woodbridge. 2004. Use of near infrared spectroscopy to predict lignin content in tropical and sub-tropical pines. **J. NIR Spectrosc.** 12: 381-390.

Horn, R.A. and V.C. Setterholm. 1990. Fiber morphology and new crops. p. 270-275. *In*: J. Janick and J.E. Simon, eds. **Advances in new crops**. Timber Press, Portland, OR.

Huang, A., G. Li, F. Fu and B. Fei. 2008. Use of visible and near infrared spectroscopy to predict klason lignin content of bamboo, Chinese fir, Paulownia, and poplar. **J. Wood Chem. Technol.** 28(3): 194-206.

Inagaki, T., M. Schwanninger, R. Kato, Y. Kurata, W. Thanapase, P. Puthson and S. Tsuchikawa. 2010. *Eucalyptus camaldulensis* density and fiber length estimated by near infrared spectroscopy. **Wood Sci. Technol.** Published online: 02 October 2010.

Jones, P.D., L.R. Schimleck, R.F. Daniels, A. Clark III and R.C. Purnell. 2008. Comparison of *Pinus taeda* L. whole-tree wood property calibrations using diffuse reflectance near infrared spectra obtained using a variety of sampling options. **Wood Sci. Technol.** 42: 385-400.

Jones, P.D., L.R. Schimleck, G.F. Peter, R.F. Daniels and A. Clark III. 2005. Nondestructive estimation of *Pinus taeda* L.tracheid morphological characteristics for samples from a wide range of sites in Georgia. **Wood Sci. Technol.** 40: 709-720.

______, ______, ______, ______ and ______. 2006. Nondestructive estimation of wood chemical composition of sections of radial wood strips by diffuse reflectance near infrared spectroscopy. **Wood Sci. Technol.** 40: 709-720.

Kelley, S.S., T.G. Rials, R. Snell, L.H. Groom and A. Sluiter. 2004. Use of near infrared spectroscopy to measure the chemical and mechanical properties of solid wood. **Wood Sci. Technol.** 38: 257-276.

Kuroda, K., N. Nishimaru, A. Izumi and D.R. Dimmel. 2002. Pyrolysis of lignin in the presence of tetramethylammonium hydroxide: A convenient method for S/G ratio determination. **J. Agric. Food Chem.** 50: 1022-1027.

Lima, C.F., L.C.A. Brabosa, C.R. Marcelo, F.O. Sivério, and J.L. Coldodette. 2008. Comparison between analytical pyrolysis and nitrobenzene oxidation for determination of syringyl/guaiacyl ratio in *Eucalytus spp*. lignin. **BioResources.** 3 (3): 701-712.

Lohninger, H. 1999. **Teach/Me—Data Analysis.** Available Source: http:// www.vias.org/tmdatanaleng/wrapnt_multivariate_data158.html, May26, 2009.

Mcclure, W.F. and S. Tsuchikawa. 2007. Instruments. *In* Y. Ozaki, W.F. McClure and A.A. Christy, eds. **Near-Infrared Spectroscopy in Food Science and Technology.** John Wiley & Sons, Inc., Hoboken, New Jersey.

Michell A.J. and L.R. Schimleck. 1998. Developing a method for the rapid assessment of pulp yield of plantation eucalypt trees beyond the year 2000. **Appita J.** 51: 428-432.

Migallon, A.J. and M.L. Marín. 2001. Pyrolysis-gas chromatography-mass spectrometry techniques for polymer degradation studies. **Encyclopedia of Chromatography** (2001): 686-689.

Muneri, A. and C.A. Raymond. 2001. Nondestructive sampling of *Eucalyptus globulus* and *E. nitens* for wood properties. – II. fiber length and coarseness. **Wood Sci. Technol.** 35: 41-56.

NSW Department of Primary Industries. n.d. **Characterisation of biochar and soil organic carbon by Py-GC-MS**. Available Source: http://www.biochar-international.org/images/Stephen_Kimber.pdf, May 24, 2009.

Nunes, C. A., C. F. Lima, L. C. A. Barbosa, J. L. Colodette, A. F. G. Gouveia and F. O. Silvério. 2010. Determination of *Eucalyptus spp.* lignin S/G ratio: A comparison between methods. **Bioresource Technol**. 101: 4056-4061.

Ohra-aho, T., J. Sipilä and T. Tamminen. 2000. Analysis of lignin and model compound by pyrolysis-GC/MS. **Sixth European Workshop on Lignocellulosic and Pulp: Advances in Lignocellulosic Chemistry towards High Quality Processes and Products, Botdeaux, France (03/09/2000).** Abstract No. INIST-CNRS, Cote INIST: Y 33672, 35400009705998.1140

Ohra-aho, T., M. Tenkanen and T. Tamminen. 2005. Direct analysis of lignin and lignin-like components from softwood kraft pulp by Py-GC/MS techniques. **J. Anal. Appl. Pyrolysis.** 74: 123-128.

Osborne, B.G. 2000. Near-infrared spectroscopy in food analysis. **Encyclopedia of Analytical Chemistry.** John Wiley & Sons, Ltd, Chichester.

Oudia, A., E. Mészáros, R. Simões, J. Queiroz and E. Jakab. 2007. Pyrolysis-GC/MS and TG/MS study of mediated laccase biodelignification of Eucalyptus globulus kraft pulp. **J. Anal. Appl. Pyrolysis.** 78: 233-242.

Ozaki, Y., S. Morita and Y. Du. 2007. Spectral analysis. *In* Y. Ozaki, W.F. McClure and A.A. Christy, eds. **Near-Infrared Spectroscopy in Food Science and Technology.** John Wiley & Sons, Inc., Hoboken, New Jersey.

Pilate, G., E. Guiney, K. Holt, M. Petit-Conil, C. Lapierre, J.-C. Leplé, B. Pollet, I. Mila, E.A. Webster, H.G. Marstorp, D.W. Hopkins, L. Jouanin, W. Boerjan, W. Schuch, D. Cornu and C. Halpin. 2002. Field and pulping performances of transgenic trees with altered lignifications. **Nat. Biotechnol.** 20: 607-612.

Poke, F.S. and C.A. Raymond. 2006. Predicting extractives, lignin and cellulose contents using near infrared spectroscopy on solid wood in *Eucalyptus globulus*. **J. Wood Chem. Technol.** 26: 187-199.

______, J.K. Wright and C.A. Raymond. 2004. Predicting extractives and lignin contents in Eucalyptus globulus using near infrared reflectance analysis. **J. Wood Chem. Technol.** 24: 55-67.

Pope, J.M. 1995. Near-infrared spectroscopy of wood products, pp. 142-151. *In* T.E. Conners and S. Banerjee, eds. **Surface Analysis of paper.** Boca Raton, CRC Press.

Ragauskas, A. 2007. **Lignin Overview.** Institute of Paper Science and Technology at Georgia Tech, Atlanta, Georgia. Available Source: http://www.ipst.gatech.edu/faculty_new/faculty_bios/raguaskas/technical_reviews/Lignin%20Overview.pdf, February 10, 2009.

Raymond, C.A., L.R. Schimleck, A. Muneri and A.J. Michell. 2001. Nondestructive sampling of *Eucalyptus globulus* and *E. nitens* for wood properties. – III. Predicted pulp yield using near infrared reflectance analysis. **Wood Sci. Technol.** 35: 203-215.

Reich, G. 2005. Near-infrared spectroscopy and imaging: Basic principles and pharmaceutical applications. **Adv. Drug. Deli. Rev.** 57: 1109-1143.

Ricardo, P. and G. Hein. 2010. Multivariate regression methods for estimation basic density in Eucalyptus wood from near infrared spectroscopic data. **Cerne.** 16: 90-96.

Ritthiruangdej, P. 2006. **Evaluation of Thai Fish Sauce Qualities by Near-Infrared Spectroscopy.** Ph.D. Thesis, Kasetsart University.

Rodrigues, J., D. Meier, O. Faix and H. Pereira. 1999. Determination of tree to tree variation in syringyl/guaiacyl ratio of *Eucalyptus globulus* wood lignin by analytical pyrolysis. **J. Anal. Appl. Pyrolysis.** 48: 121-128.

Ruiz, J., J. Rodriguez, J. Baeza and J. Freer. 2005. Estimating density and pulping yield of E. globulus wood: comparison of near-infrared (NIR) and mid-infrared (MIR). **J. Chil. Chem. Soc.** 50: 565-568.

Sakakibara, A. and Y. Sono. 2001. Chapter 4: Chemistry of lignin, pp. 109-173. *In* D.N.-S. Hon and N. Shiraishi, eds. **Wood and Cellulosic Chemistry.** Marcel Dkker, New York.

Schimleck, L.R. n.d. **Near infrared spectroscopy: a rapid, non-destructive method for measuring wood properties and its application to tree breeding.** Available Source: http://proceedings.com.au/afgc/papers%20(pdf)/weds%20schimleck.pdf, May 26, 2009.

______ and R. Evans. 2003. Estimation of air-dry density of increment cores by near infrared spectroscopy. **Appita J.** 56: 312-317.

______ and R. Evans. 2004. Estimation of *Pinus radiate* D. Don tracheid morphological characteristics by near infrared spectroscopy. **Holzforschung.** 58: 66-73.

______ and J. French. 2002. Application of NIR spectroscopy to clonal *Eucalyptus globulus* samples covering a narrow range of pulp yield. **Appita J.** 55: 149-154.

______, A.J. Michell, C.A. Raymond amd A. Muneri. 1999. Estimation of basic density of *Eucalyptus globulus* using near-infrared spectroscopy. **Can. J. For. Res.** 29(2): 194-201.

______, P.D. Kube and C.A. Raymond. 2004. Genetic improvement of kraft pulp yield in *Eucalyptus nitens* using cellulose content determined by near infrared spectroscopy. **Can. J. For. Res.** 34: 2363-2370.

______, ______ and ______. 2005a. Estimation of whole-tree kraft pulp yield of *Eucalyptus nitrens* using near-infrared spectra collected from increment cores. **Can. J. For. Res.** 35: 2797-2805.

______, P. Payne and R.H. Wearne. 2005b. Determination of important pulp properties of hybrid poplar by near infrared spectroscopy. **Wood Fiber Sci.** 37: 462-471.

Seca, A.M.L. da and e F.M. de J. Domingues. 2006. Basic density and pulp yield relationship with some chemical parameters in eucalyptus trees. **Pesq. Agropec. Bras.** 41: 1687-1691.

Siesler, H.W., Y. Ozaki, S. Kawata and H.M. Heise. 2002. **Near infrared spectroscopy: principles, instruments and applications**. Wiley VCH, Germany.

Sitholé, B.B. 2000. Pyrolysis in the pulp and paper industry. **Encyclopedia of Analytical Chemistry.** John Wiley & Sons, Ltd.

Sonoda,T., T. Ona, H. Yokoi, Y. Ishida, H. Ohtani and S. Tsuge. 2001. Quantitative analysis of detailed lignin monomer composition by pyrolysis-gas chromatography combined with preliminary acetylation of the samples. **Anal. Chem.** 73: 5429-5435.

Suttiwijitpukdee, N. 2006. **Determination of Chemical Compositions and Pulp Yield of Eucalyptus camaldulensis Wood by Near Infrared Spectroscopy.** M.S. Thesis, Kasetsart University. *Cited* M.K. Boysworth and K.S. Booksh. 2001. Aspects of multivariate calibration applied to near-infrared spectroscopy. *In* D.A. Burns and E.W. Ciurczak. **Handbook of Near-Infrared Analysis.** Marcel Dekker, New York.

Technical Association of the Pulp and Paper Industry. 2002a. **Coarseness of pulp fibers.** TAPPI T 234-cm02.

______. 2002b. **Basic density and moisture content of pulpwood.** TAPPI T 258-om02.

Terdwongworakul, A., V. Punsuwan, W. Thanapase and S. Tsuchikawa. 2005. Rapid assessment of wood chemical properties and pulp yield of *Eucalyptus camaldulensis* in Thailand tree plantations by near infrared spectroscopy for improving wood selection for high quality pulp. **J. Wood. Sci.** 51: 167-171.

Tsuchikawa, S. 2007. A review of recent near infrared research for wood and paper. **Applied Spectroscopy Reviews.** 42: 1, 43-7

Tsutsumi, Y., R. Kondo, K. Sakai and H. Imamura. 1995. The difference of reactivity between syringyl lignin and guaiacyl lignin in alkaline systems. **Holzforschung.** 49: 423-428.

Via, B.K., C.-L. So, T.F. Shupe, M. Stine and L.H. Groom. 2005. Ability of near infrared spectroscopy to monitor air-dry density distribution and variation of wood. **Wood and Fiber Science**. 37(3):394-402.

Weantong, S. 2005. **Evaluation of Plam Oil Quality by Fourier-Transform Near Infrared Spectroscopy.** M.S. Thesis, Kasetsart University. (inThai) *Cited* S. Kawano and S. Saranwong. 2004. NIR instruments and sample presentation, pp.10-19. **The Second Workshop on NIR Spectroscopy in Thailand.**

Wolf, C.J., M.A. Grayson and D.L. Fanter. 1980. Pyrolysis gas chromatography of polymer: Pyrolysis make possible the GC characterization of complex intractable materials. **Anal. Chem.** 53 (3): 348-358.

Workman, J. Jr. and L. Weyer. 2008. **Practical guide to interpretive near-infrared spectroscopy**. CRC Press, New York.

Wright, J.A., M.D. Birkett and M.J.T. Gambino. 1990. Prediction of pulp yield and cellulose content from wood samples using near-infrared reflectance spectroscopy. **Tappi J.** 73 (8): 164-166.

Yamada, T., T.-F. Yen, H.-M. Chang, L. Li, J.F. Kadla and V.L. Chiang. 2006. Rapid analysis of transgenic trees using transmittance near-infrared spectroscopy (NIR). **Holzforschung.** 60: 24-28.

Yen, T.-F., H.-M. Chang and J.F. Kadla. 2004. Rapid prediction of solid wood lignin content using transmittance near-infrared spectroscopy. **J. Agric. Food Chem.** 52: 1435-1439.

Yokoi, H., Y. Ishida, H. Ohtani, S. Tsuge, T. Sonoda and T. Ona. 1999. Characterization of within-tree variation of lignin components in *Eucalyptus camaldulensis* by pyrolysis-gas chromatography. **Analyst.** 124: 669-674.

______, T. Nakase, Y. Ishida, H. Ohtani, S. Tsuge, T. Sonoda and T. Ona. 2001. Discriminative analysis of *Eucalyptus calmaldulensis* grown from seeds of origins based on lignin components measured by pyrolysis-gas chromatography. **J. Anal. Appl. Pyrolysis.** 57: 145-152.

APENDICES

Appendix A

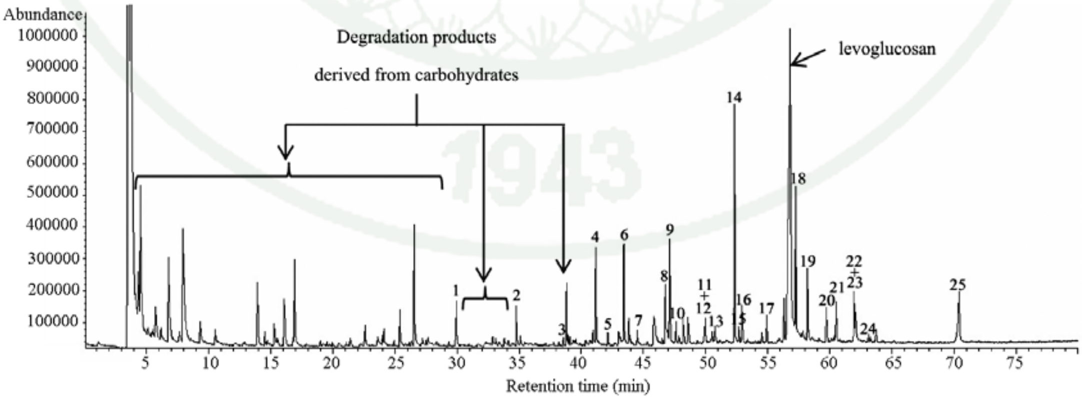
Analysis of S/G ratio of lignin by using Py-GC/MS

The pyrolysis system consisted of a Frontier Lab Micro furnace Double-shot Pyrolyzer (Py-2020iD) equipped with an Autosampler (AS-1020E). Each sample (~80 μg) was pyrolyzed at 500˚C for 10 s. The GC interface was kept at 360˚C and was purged with a continuous helium flow for 15 s prior to pyrolysis. After pyrolysis the products were transferred (split 1:50) into the capillary column (DB1701; 60 m × 0.25 mm × 0.25 μm film) via directly connected GC-injector. The gas chromatography (Agilent 6890) was performed with helium as the carrier gas at a flow rate of 1 ml/min. The oven programme started off with 4 min of isothermal condition at 45˚C, then was raised at a rate of 3˚C/min to 280˚C and was finally kept at this temperature for 20 min. For mass spectral detection an Agilent 5973N MSD was used with and electron impact ionization energy of 70 *eV*. The scan range was 35-400 *m/z*. Triplicate analyses were performed and averaged for display in the tables and illustrations. Identification of the products was carried out by library search (NIST02) and comparison with literature source (Bremer, 1991; Lima *et al.*, 2008; Nunes *et al.*, 2010; Yokoi *et al.*, 1999).

For determination of S/G ratio with MS signal (TIC), the extracted single ion peak areas (EIC) of all lignin derived products in the pyrogram were integrated, and then convert of each EIC peak area to theoretical TIC peak area again with conversion factor which is calculated by the shared amount of the selected ion abundance of the lignin product. This calculation way could be solved the problem about overlapping peaks and also allowed to use all identifiable lignin products for calculation the S/G value. Moreover, the S/G values were not overestimation, the value seem to close with the value from standard wet chemical S/G ratio analysis, such as nitrobenzene oxidation (Lima *et al.*, 2008; Nunes *et al.*, 2010) or thioacidolysis (Yokoi *et al.*, 1999) methods.

The Py-GC/MS instrument for determination of wood lignin compositions used in this experiment is shown in Appendix Figure A1. Appendix Figure A2 shows a typical pyrogram of a Eucalyptus wood meal at 500˚C observed by TIC signal of MS. Many pyrolyzates derived from lignin were observed as peaks 1-25 on the pyrogram along with those derived from cellulose and hemicelluloses such as levoglucosan. Appendix Table A1 shows the assignment of these characteristic peaks from lignin and calculated their relative percent amounts by used theoretical TIC signal of the MS detection. The 25 lignin peaks were assigned to 13 S and 12 G pyrolyzates with the corresponding side-chain structures. The S/G ratio for each sample can be calculated by sum of the peak areas of characteristic peaks of S unit divided with sum of the peak areas of characteristic peaks of G unit.

Appendix Figure A1 Py-GC/MS instrument for determination of S/G ratio of lignin



Appendix Figure A2 Total ion chromatogram (TIC) of a sample of Eucalyptus wood. Peak identities for the lignin-derived compounds are shown in Appendix Table C1

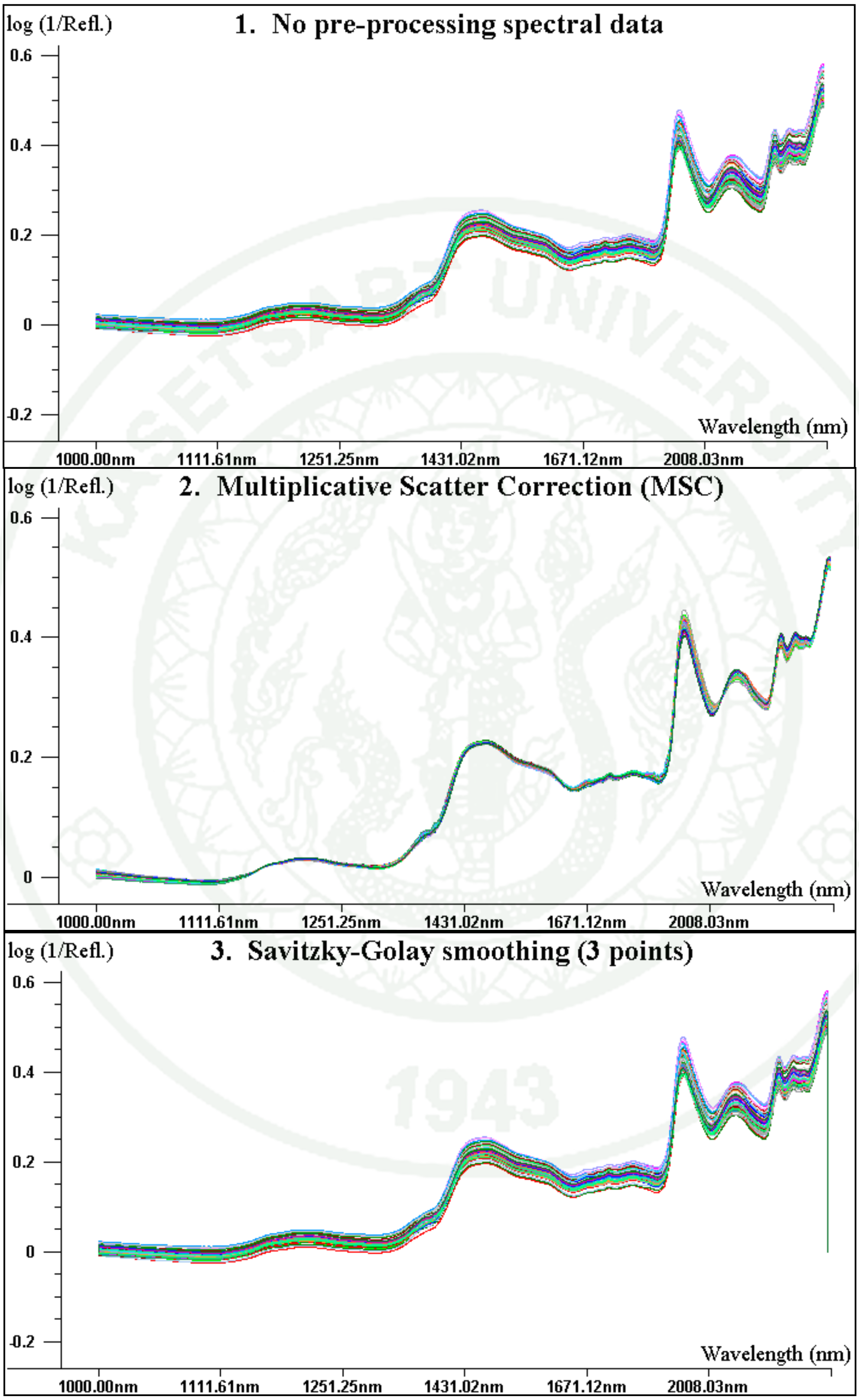
Appendix Table A1 Peak assignment of lignin derived products in the pyrogram

Peak No.	Lignin derivative marker	M_w	Origin	Ion key	Relative amount* (%)
1	Guaiacol	124	G	m/z 109	2.52
2	4-Methylguaiacol	138	G	m/z 138	3.43
3	4-Ethylguaiacol	152	G	m/z 137	0.55
4	Vinyl guaiacol	150	G	m/z 150	12.52
5	Eugenol	164	G	m/z 164	1.13
6	Syringol	154	S	m/z 154	7.33
7	*cis*-Isoeugenol	164	G	m/z 164	0.87
8	*trans*-isoeugono	164	G	m/z 164	4.85
9	4-Methylsyringol	168	S	m/z 168	6.70
10	Vanillin	152	G	m/z 151	1.95
11	Homovanillin	166	G	m/z 137	1.40
12	4-Ethylsyringol	182	S	m/z 167	1.08
13	Acetoguaiacone	166	G	m/z 151	1.06
14	4-Vinylsyringol	180	S	m/z 180	15.08
15	Guaiacylacetone	180	G	m/z 137	1.47
16	4-Allylsyringol	194	S	m/z 194	2.30
17	*cis*-4-Propenylsyringol	194	S	m/z 194	1.63
18	*trans*-4-Propenylsyringol	194	S	m/z 194	8.75
19	Syringaldehyde	182	S	m/z 182	5.60
20	Homosyringaldehyde	196	S	m/z 167	2.46
21	Acetosyringone	196	S	m/z 181	3.01
22	Syringylacetone	210	S	m/z 167	3.09
23	Coniferaldehyde	178	G	m/z 178	1.67
24	Propiosyringone	210	S	m/z 181	0.40
25	Sinapaldehyde	208	S	m/z 208	9.15

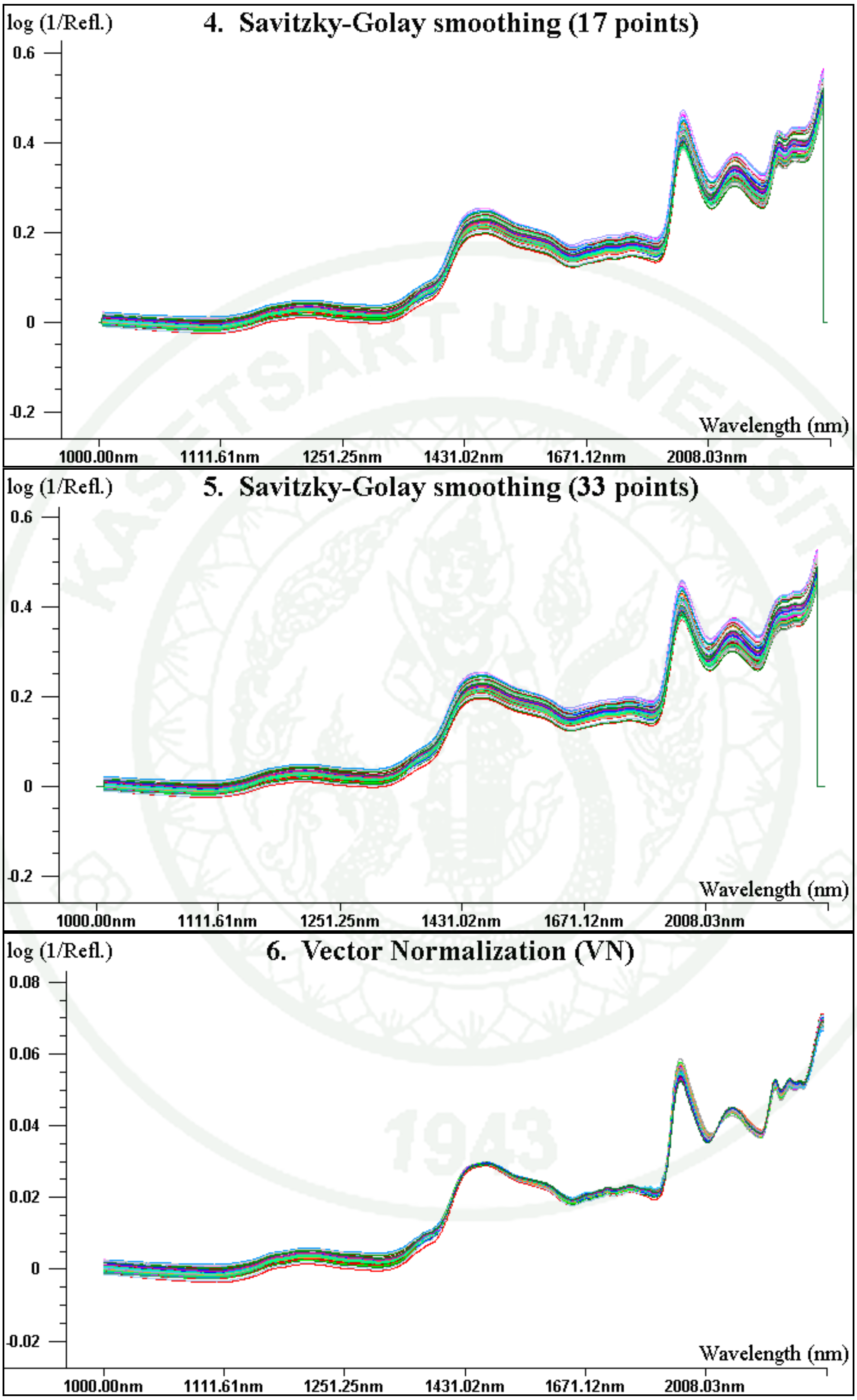
Remark: * relative amount in percent was calculated after conversion EIC to theoretical TIC peak area already

Appendix B

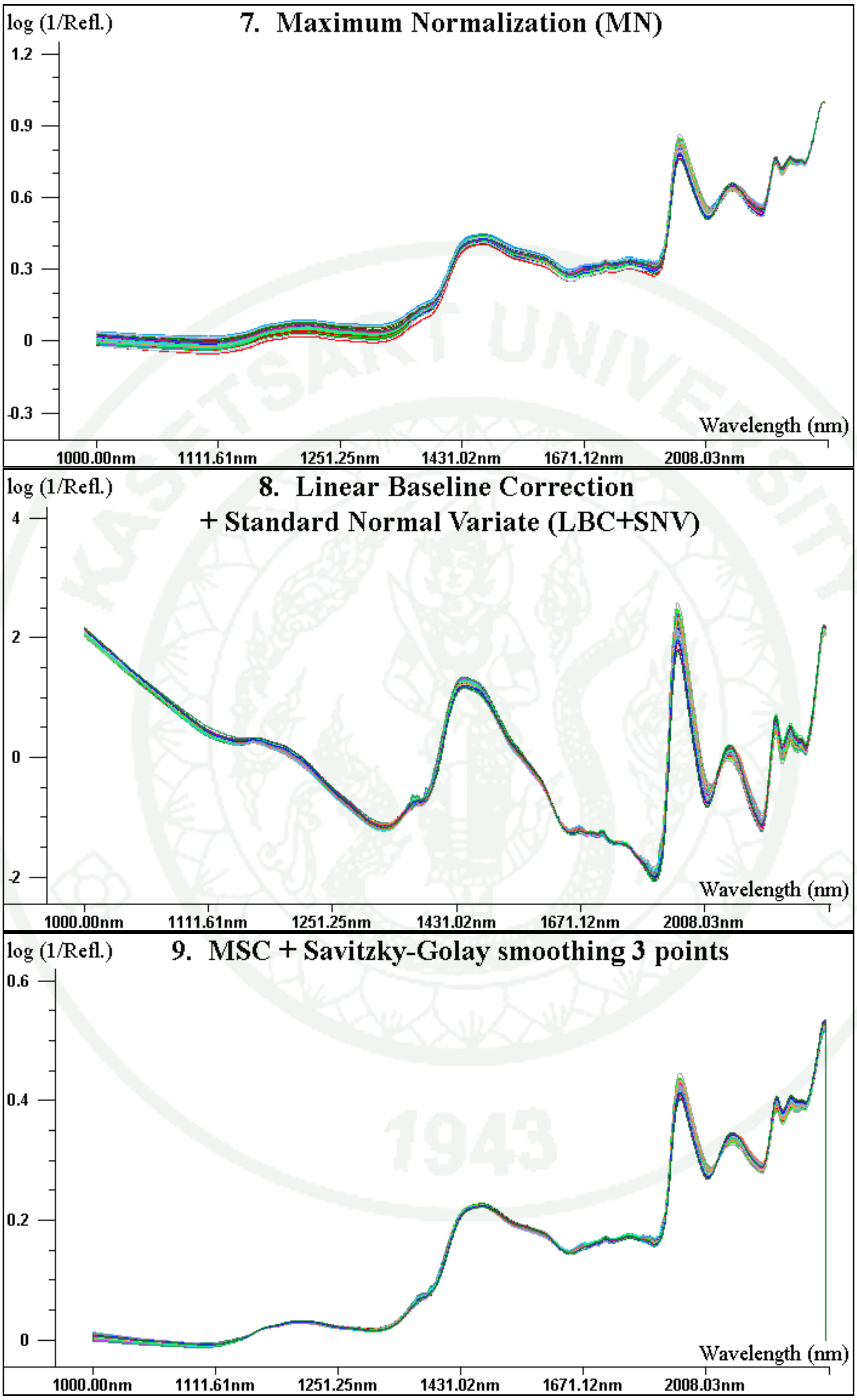
Mathematic pre-processing methods for NIR spectral data.



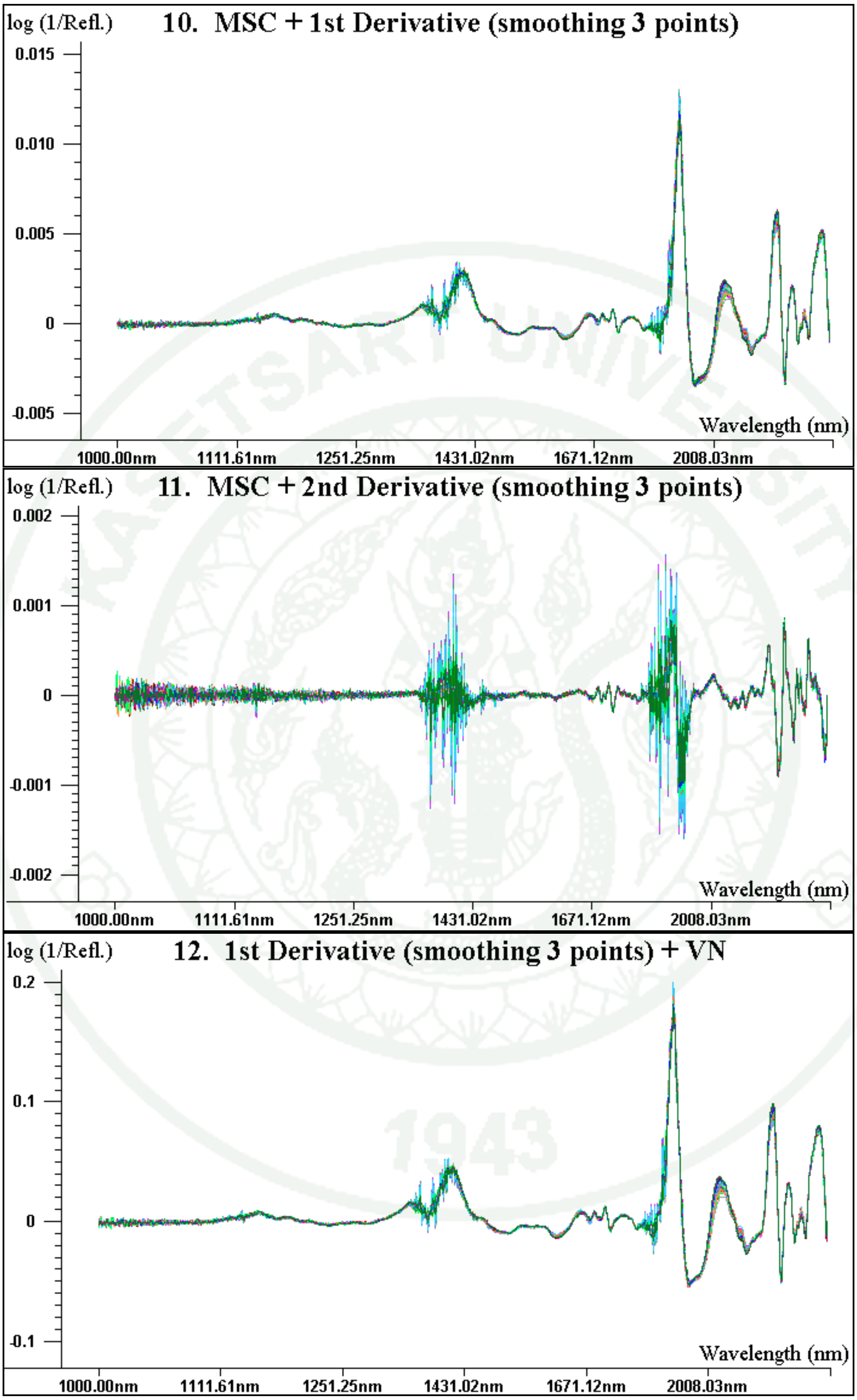
Appendix Figure B1 Various mathematic pre-processing methods for NIR spectra



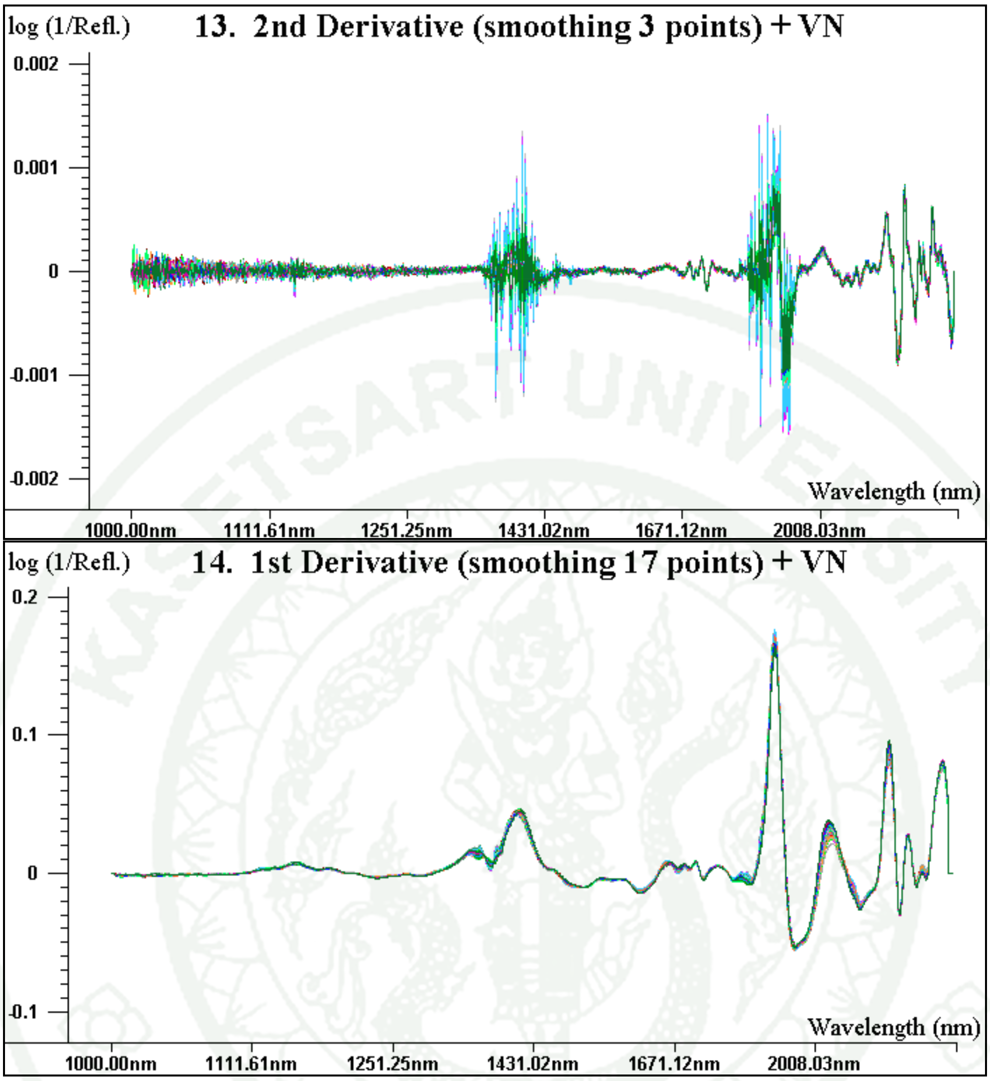
Appendix Figure B1 (Continued)



Appendix Figure B1 (Continued)



Appendix Figure B1 (Continued)



Appendix Figure B1 (Continued)

Appendix C

Multivariate data analysis results from the experiment

Appendix Table C1 Statistical results of wood basic density NIR model in range 1000-2500 nm obtained by PCR with different spectral pre-processing methods

No.	NIR spectral pre-processing	No. of PC	Calibration set (N=45)					Validation set (N=15)				
			R^2	RMSEC	SEC	Bias	RPD_c	Adjusted R^2	RMSEP	SEP	Bias	RPD_p
1	No pre-processing	13	0.8945	16.7734	16.963	6.78E-07	3.08	0.8810	17.7046	18.321	-4.27E-01	2.85
2	MSC	14	0.8624	19.1534	19.370	5.83E-05	2.70	0.8492	19.9295	20.625	3.76E-01	2.53
3	SG smoothing 3 points	13	0.8942	16.7952	16.985	3.59E-05	3.07	0.8807	17.7263	18.343	-4.42E-01	2.85
4	SG smoothing 17 points	13	0.8927	16.9141	17.105	-1.29E-05	3.05	0.8783	17.9028	18.526	-4.42E-01	2.82
5	SG smoothing 33 points	13	0.8918	16.9855	17.177	1.15E-05	3.04	0.8780	17.9291	18.554	3.75E-01	2.81
6	VN	14	0.8716	18.5044	18.714	6.99E-05	2.79	0.8306	21.1226	21.854	-6.30E-01	2.39
7	MN	8	0.8027	22.9383	23.197	-3.84E-04	2.25	0.7795	24.1018	24.434	-4.86	2.14
8	LBC + SNV	12	0.8522	19.8485	20.073	2.24E-05	2.60	0.7154	27.3775	28.338	1.62E-01	1.84
9	MSC+SG smoothing 3pts	14	0.8620	19.1790	19.396	2.71E-05	2.69	0.8477	20.0282	20.725	5.02E-01	2.52
10	MSC + 1st Der (3 points)	7	0.7533	25.6481	25.938	2.71E-06	2.01	0.8731	18.2820	17.875	6.00	2.92
11	MSC + 2nd Der (3 points)	5	0.6732	29.5177	29.851	6.78E-06	1.75	0.8623	19.0464	16.594	1.03E+01	3.15
12	1st Der (3 points) + VN	7	0.7601	25.2922	25.578	8.14E-06	2.04	0.8847	17.4256	16.971	5.90	3.08
13	2nd Der (3 points) + VN	4	0.5385	35.0774	35.474	6.78E-06	1.47	0.8387	20.6095	19.107	9.17	2.73
14	1st Der (17 points) + VN	8	0.7573	25.4372	25.725	5.43E-06	2.03	0.8835	17.5157	17.992	-2.16	2.90

Appendix Table C2 Statistical results of wood basic density NIR model in range 1000-2500 nm obtained by PLSR with different spectral pre-processing methods

No.	NIR spectral pre-processing	No. of PC	Calibration set (N=45)					Validation set (N=15)				
			R^2	RMSEC	SEC	Bias	RPD_c	Adjusted R^2	RMSEP	SEP	Bias	RPD_p
1	No pre-processing	11	0.9217	14.4531	14.616	1.56E-05	3.57	0.9013	16.1242	16.659	9.92E-01	3.13
2	MSC	11	0.9394	12.7068	12.850	-5.09E-05	4.06	0.8811	17.6960	18.051	3.01	2.89
3	SG smoothing 3 points	11	0.9202	14.5903	14.755	2.92E-05	3.54	0.9008	16.1631	16.706	8.78E-01	3.13
4	SG smoothing 17 points	11	0.9112	15.3915	15.565	6.78E-06	3.35	0.8913	16.9198	17.511	2.83E-01	2.98
5	SG smoothing 33 points	12	0.9187	14.7220	14.888	1.76E-05	3.51	0.8958	16.5674	16.752	3.54	3.12
6	VN	9	0.8761	18.1770	18.382	4.27E-05	2.84	0.8074	22.5235	23.281	1.19	2.24
7	MN	12	0.9352	13.1408	13.289	4.19E-03	3.93	0.8316	21.0621	21.733	-1.67	2.40
8	LBC + SNV	5	0.8079	22.6318	22.888	1.09E-05	2.28	0.6785	29.0984	30.020	-2.37	1.74
9	MSC+SG smoothing 3pts	11	0.9374	12.9156	13.062	-2.04E-05	4.00	0.8817	17.6548	17.984	3.14	2.90
10	MSC + 1st Der (3 points)	3	0.8308	21.2413	21.481	1.09E-05	2.43	0.8665	18.7491	18.705	5.00	2.79
11	MSC + 2nd Der (3 points)	2	0.7576	25.4234	25.711	5.43E-06	2.03	0.8716	18.3917	16.018	9.94	3.26
12	1st Der (3 points) + VN	3	0.8316	21.1875	21.427	1.63E-05	2.44	0.8852	17.3905	17.222	5.06	3.03
13	2nd Der (3 points) + VN	2	0.6791	29.2522	29.583	1.49E-05	1.77	0.8399	20.5372	17.870	11.12	2.92
14	1st Der (17 points) + VN	3	0.7725	24.6303	24.909	6.78E-06	2.10	0.8793	17.8303	18.015	-3.88	2.90

Appendix Table C3 Statistical results of wood basic density NIR model in range 1400-2500 nm obtained by PCR with different spectral pre-processing methods

No.	NIR spectral pre-processing	No. of PC	Calibration set (N=45)					Validation set (N=15)				
			R^2	RMSEC	SEC	Bias	RPD_c	Adjusted R^2	RMSEP	SEP	Bias	RPD_p
1	No pre-processing	14	0.8902	17.1085	17.302	8.82E-06	3.02	0.8686	18.6050	19.250	-5.20E-01	2.71
2	MSC	12	0.8368	20.8584	21.094	1.32E-04	2.48	0.8013	22.8770	23.665	8.17E-01	2.21
3	SG smoothing 3 points	14	0.8900	17.1290	17.323	4.41E-05	3.01	0.8683	18.6232	19.269	-5.16E-01	2.71
4	SG smoothing 17 points	14	0.8885	17.2454	17.440	-2.44E-05	2.99	0.8661	18.7828	19.430	-6.57E-01	2.69
5	SG smoothing 33 points	13	0.8874	17.3232	17.519	3.39E-06	2.98	0.8611	19.1260	19.792	-4.42E-01	2.64
6	VN	14	0.8930	16.8909	17.082	4.75E-05	3.06	0.8190	21.8349	21.914	5.34	2.38
7	MN	22	0.9237	14.2630	14.424	4.15E-03	3.62	0.7903	23.5011	24.315	6.96E-01	2.15
8	LBC + SNV	11	0.8381	20.7794	21.014	1.97E-05	2.48	0.7100	27.6401	28.610	1.32E-01	1.83
9	MSC+SG smoothing 3pts	11	0.8368	20.8617	21.097	-3.87E-05	2.48	0.7944	23.2728	24.072	8.97E-01	2.17
10	MSC + 1st Der (3 points)	6	0.7235	27.1534	27.460	1.83E-05	1.90	0.8634	18.9699	19.628	-5.34E-01	2.66
11	MSC + 2nd Der (3 points)	6	0.6932	28.5990	28.922	9.49E-06	1.81	0.8621	19.0566	17.841	8.13	2.93
12	1st Der (3 points) + VN	6	0.7437	26.1390	26.434	5.43E-06	1.98	0.8393	20.5741	21.163	2.30	2.47
13	2nd Der (3 points) + VN	6	0.5516	34.5769	34.968	1.09E-05	1.49	0.8544	19.5802	17.812	9.34	2.93
14	1st Der (17 points) + VN	7	0.7206	27.2926	27.601	9.49E-06	1.89	0.8213	21.6969	22.368	1.95	2.33

Appendix Table C4 Statistical results of wood basic density NIR model in range 1400-2500 nm obtained by PLSR with different spectral pre-processing methods

No.	NIR spectral pre-processing	No. of PC	Calibration set (N=45)					Validation set (N=15)				
			R^2	RMSEC	SEC	Bias	RPD_c	Adjusted R^2	RMSEP	SEP	Bias	RPD_p
1	No pre-processing	9	0.8967	16.5962	16.784	1.63E-05	3.11	0.8568	19.4188	20.095	-4.31E-01	2.60
2	MSC	8	0.8907	17.0681	17.261	4.07E-05	3.03	0.8075	22.5189	23.112	2.92	2.26
3	SG smoothing 3 points	10	0.9109	15.4128	15.587	3.53E-05	3.35	0.8672	18.7021	19.292	1.55	2.71
4	SG smoothing 17 points	10	0.9027	16.1037	16.286	1.36E-05	3.21	0.8731	18.2854	18.905	8.83E-01	2.76
5	SG smoothing 33 points	10	0.8970	16.5740	16.761	-2.65E-05	3.12	0.8703	18.4833	19.127	4.42E-01	2.73
6	VN	9	0.8824	17.7093	17.909	3.39E-06	2.92	0.8279	21.2911	21.897	2.40	2.38
7	MN	11	0.9250	14.1445	14.304	3.62E-03	3.65	0.7846	23.8212	24.587	1.79	2.12
8	LBC + SNV	8	0.8633	19.0893	19.305	1.36E-05	2.70	0.7076	27.7531	28.715	-8.02E-01	1.82
9	MSC+SG smoothing 3pts	8	0.8908	17.0666	17.259	-7.73E-05	3.03	0.8054	22.6406	23.229	3.00	2.25
10	MSC + 1st Der (3 points)	3	0.8029	22.9247	23.184	1.63E-05	2.25	0.9068	15.6681	16.197	-8.06E-01	3.22
11	MSC + 2nd Der (3 points)	2	0.8023	22.9610	23.220	6.78E-06	2.25	0.8450	20.2083	18.717	9.02	2.79
12	1st Der (3 points) + VN	3	0.7957	23.3409	23.605	1.09E-05	2.21	0.8568	19.4218	20.084	8.53E-01	2.60
13	2nd Der (3 points) + VN	3	0.7434	26.1570	26.453	8.14E-06	1.97	0.8392	20.5811	17.663	1.15E+01	2.96
14	1st Der (17 points) + VN	3	0.7159	27.5245	27.836	6.78E-06	1.88	0.7931	23.3434	23.789	-4.09	2.20

Appendix Table C5 Statistical results of syringyl/guaiacyl ratio NIR model in range 1000-2500 nm obtained by PCR with different spectral pre-processing methods

No.	NIR spectral pre-processing	No. of PC	Calibration set (N=45)					Validation set (N=15)				
			R^2	RMSEC	SEC	Bias	RPD_c	Adjusted R^2	RMSEP	SEP	Bias	RPD_p
1	No pre-processing	11	0.8617	0.1267	0.1281	7.95E-09	2.69	0.8603	0.1291	0.1317	2.21E-02	2.62
2	MSC	14	0.9391	0.0841	0.0850	1.19E-07	4.05	0.8747	0.1223	0.1266	-1.08E-03	2.72
3	SG smoothing 3 points	11	0.8623	0.1264	0.1278	-3.18E-08	2.70	0.8622	0.1283	0.1308	2.19E-02	2.63
4	SG smoothing 17 points	11	0.8648	0.1252	0.1266	-1.06E-07	2.72	0.8732	0.1230	0.1255	2.08E-02	2.74
5	SG smoothing 33 points	11	0.8649	0.1252	0.1266	3.26E-07	2.72	0.8806	0.1194	0.1218	2.03E-02	2.83
6	VN	11	0.8747	0.1205	0.1219	9.27E-08	2.83	0.8716	0.1238	0.1269	1.71E-02	2.71
7	MN	11	0.8455	0.1339	0.1354	1.23E-06	2.54	0.8223	0.1457	0.1495	1.90E-02	2.30
8	LBC + SNV	12	0.9420	0.0820	0.0829	-1.03E-07	4.15	0.8585	0.1300	0.1342	8.71E-03	2.57
9	MSC+SG smoothing 3pts	14	0.9392	0.0840	0.0849	-3.15E-07	4.05	0.8745	0.1224	0.1267	-1.60E-03	2.72
10	MSC + 1st Der (3 points)	10	0.8869	0.1146	0.1159	3.97E-08	2.97	0.7755	0.1637	0.1684	1.83E-02	2.05
11	MSC + 2nd Der (3 points)	10	0.8351	0.1383	0.1399	3.71E-08	2.46	0.7443	0.1747	0.1766	3.79E-02	1.95
12	1st Der (3 points) + VN	10	0.8830	0.1165	0.1178	5.56E-08	2.92	0.7753	0.1638	0.1680	2.19E-02	2.05
13	2nd Der (3 points) + VN	6	0.6255	0.2084	0.2108	4.50E-08	1.63	0.6023	0.2179	0.2186	5.36E-02	1.58
14	1st Der (17 points) + VN	2	0.7473	0.1732	0.1732	5.30E-08	1.99	0.8193	0.1469	0.1449	4.43E-02	2.38

Appendix Table C6 Statistical results of syringyl/guaiacyl ratio NIR model in range 1000-2500 nm obtained by PLSR with different spectral pre-processing methods

No.	NIR spectral pre-processing	No. of PC	Calibration set (N=45)					Validation set (N=15)				
			R^2	RMSEC	SEC	Bias	RPD_c	Adjusted R^2	RMSEP	SEP	Bias	RPD_p
1	No pre-processing	9	0.9109	0.1017	0.1028	1.35E-07	3.35	0.8633	0.1278	0.1287	2.93E-02	2.68
2	MSC	8	0.9138	0.1000	0.1011	1.75E-07	3.41	0.8739	0.1227	0.1252	2.06E-02	2.75
3	SG smoothing 3 points	9	0.9108	0.1017	0.1029	-2.09E-07	3.35	0.8634	0.1277	0.1286	2.96E-02	2.68
4	SG smoothing 17 points	10	0.9369	0.0856	0.0865	-1.35E-07	3.98	0.8689	0.1251	0.1295	1.22E-03	2.66
5	SG smoothing 33 points	10	0.9328	0.0883	0.0893	5.43E-07	3.86	0.8724	0.1234	0.1278	1.58E-03	2.70
6	VN	8	0.9123	0.1009	0.1020	-8.21E-08	3.38	0.8741	0.1226	0.1269	-5.07E-04	2.71
7	MN	5	0.8529	0.1306	0.1321	3.23E-07	2.61	0.8476	0.1349	0.1376	2.29E-02	2.50
8	LBC + SNV	11	0.9743	0.0546	0.0553	-1.40E-07	6.23	0.8698	0.1247	0.1290	3.68E-03	2.67
9	MSC+SG smoothing 3pts	8	0.9132	0.1003	0.1015	-3.76E-07	3.39	0.8724	0.1234	0.1259	2.09E-02	2.74
10	MSC + 1st Der (3 points)	6	0.9853	0.0413	0.0418	5.30E-08	8.24	0.8152	0.1485	0.1537	2.11E-03	2.24
11	MSC + 2nd Der (3 points)	7	0.9969	0.0188	0.0190	2.38E-08	18.1	0.7963	0.1559	0.1589	2.72E-02	2.17
12	1st Der (3 points) + VN	6	0.9863	0.0398	0.0402	2.65E-08	8.56	0.8247	0.1446	0.1496	5.41E-03	2.30
13	2nd Der (3 points) + VN	3	0.8758	0.1200	0.1214	6.62E-08	2.84	0.6979	0.1899	0.1931	3.55E-02	1.78
14	1st Der (17 points) + VN	6	0.9741	0.0548	0.0554	1.33E-08	6.22	0.8194	0.1468	0.1519	3.93E-03	2.27

Appendix Table C7 Statistical results of syringyl/guaiacyl ratio NIR model in range 1400-2500 nm obtained by PCR with different spectral pre-processing methods

No.	NIR spectral pre-processing	No. of PC	Calibration set (N=45)					Validation set (N=15)				
			R^2	RMSEC	SEC	Bias	RPD_c	Adjusted R^2	RMSEP	SEP	Bias	RPD_p
1	No pre-processing	9	0.8823	0.1169	0.1182	1.75E-07	2.91	0.8735	0.1229	0.1244	2.57E-02	2.77
2	MSC	8	0.8664	0.1245	0.1259	2.04E-07	2.74	0.8207	0.1463	0.1493	2.47E-02	2.31
3	SG smoothing 3 points	9	0.8830	0.1165	0.1178	6.62E-08	2.92	0.8740	0.1227	0.1242	2.56E-02	2.77
4	SG smoothing 17 points	9	0.8878	0.1141	0.1154	-2.36E-07	2.99	0.8744	0.1224	0.1241	2.48E-02	2.78
5	SG smoothing 33 points	9	0.8891	0.1134	0.1147	4.95E-07	3.00	0.8706	0.1243	0.1259	2.53E-02	2.74
6	VN	8	0.8422	0.1353	0.1368	1.33E-07	2.52	0.8719	0.1237	0.1190	4.56E-02	2.89
7	MN	12	0.8866	0.1147	0.1160	-7.15E-07	2.97	0.8403	0.1381	0.1411	2.22E-02	2.44
8	LBC + SNV	11	0.9415	0.0824	0.0833	-1.09E-07	4.13	0.8494	0.1341	0.1388	1.83E-03	2.48
9	MSC+SG smoothing 3pts	8	0.8665	0.1245	0.1259	-5.25E-07	2.74	0.8211	0.1461	0.1491	2.45E-02	2.31
10	MSC + 1st Der (3 points)	9	0.8734	0.1212	0.1226	5.56E-08	2.81	0.7227	0.1819	0.1836	4.05E-02	1.88
11	MSC + 2nd Der (3 points)	7	0.8475	0.1330	0.1345	2.65E-08	2.56	0.7401	0.1761	0.1814	1.79E-02	1.90
12	1st Der (3 points) + VN	8	0.8764	0.1198	0.1211	4.24E-08	2.84	0.7424	0.1754	0.1786	3.15E-02	1.93
13	2nd Der (3 points) + VN	10	0.8213	0.1440	0.1456	4.50E-08	2.37	0.6831	0.1945	0.1984	3.30E-02	1.74
14	1st Der (17 points) + VN	7	0.8754	0.1202	0.1216	3.44E-08	2.83	0.7320	0.1789	0.1812	3.69E-02	1.90

Appendix Table C8 Statistical results of syringyl/guaiacyl ratio NIR model in range 1400-2500 nm obtained by PLSR with different spectral pre-processing methods

No.	NIR spectral pre-processing	No. of PC	Calibration set (N=45)					Validation set (N=15)				
			R^2	RMSEC	SEC	Bias	RPD_c	Adjusted R^2	RMSEP	SEP	Bias	RPD_p
1	No pre-processing	8	0.9265	0.0924	0.0934	2.65E-07	3.69	0.8455	0.1358	0.1389	2.10E-02	2.48
2	MSC	7	0.9199	0.0964	0.0975	1.25E-07	3.53	0.8485	0.1345	0.1392	3.71E-03	2.48
3	SG smoothing 3 points	8	0.9264	0.0924	0.0934	-2.09E-07	3.69	0.8460	0.1356	0.1387	2.11E-02	2.48
4	SG smoothing 17 points	8	0.9246	0.0936	0.0946	-1.72E-07	3.64	0.8497	0.1339	0.1368	2.16E-02	2.52
5	SG smoothing 33 points	8	0.9202	0.0962	0.0973	5.96E-07	3.54	0.8526	0.1327	0.1353	2.27E-02	2.55
6	VN	7	0.9009	0.1072	0.1085	-9.80E-08	3.18	0.8893	0.1149	0.1178	1.58E-02	2.92
7	MN	5	0.8554	0.1295	0.1310	-1.96E-07	2.63	0.8322	0.1415	0.1323	6.07E-02	2.60
8	LBC + SNV	10	0.9694	0.0596	0.0603	-1.33E-07	5.71	0.8637	0.1276	0.1320	2.71E-03	2.61
9	MSC+SG smoothing 3pts	7	0.9199	0.0964	0.0975	-5.59E-07	3.53	0.8485	0.1345	0.1391	3.88E-03	2.48
10	MSC + 1st Der (3 points)	5	0.9729	0.0561	0.0567	6.09E-08	6.07	0.7807	0.1618	0.1675	-2.87E-03	2.06
11	MSC + 2nd Der (3 points)	6	0.9952	0.0237	0.0240	5.30E-08	14.4	0.7613	0.1688	0.1741	1.48E-02	1.98
12	1st Der (3 points) + VN	5	0.9724	0.0566	0.0572	1.85E-08	6.02	0.7905	0.1582	0.1637	-2.99E-03	2.10
13	2nd Der (3 points) + VN	4	0.9461	0.0791	0.0800	2.65E-08	4.31	0.6680	0.1991	0.2061	1.13E-03	1.67
14	1st Der (17 points) + VN	5	0.9555	0.0718	0.0727	1.33E-08	4.74	0.8166	0.1480	0.1529	9.02E-03	2.25

Appendix Table C9 Statistical results of fiber length NIR model in range 1000-2500 nm obtained by PCR with different spectral pre-processing methods

No.	NIR spectral pre-processing	No. of PC	Calibration set (N=45)					Validation set (N=15)				
			R^2	RMSEC	SEC	Bias	RPD_c	Adjusted R^2	RMSEP	SEP	Bias	RPD_p
1	No pre-processing	6	0.3893	0.0431	0.0436	4.37E-08	1.28	0.6736	0.0285	0.0271	1.13E-02	2.06
2	MSC	5	0.3720	0.0437	0.0442	4.11E-08	1.26	0.6322	0.0303	0.0295	1.02E-02	1.89
3	SG smoothing 3 points	6	0.3893	0.0431	0.0436	2.78E-08	1.28	0.6736	0.0285	0.0271	1.13E-02	2.06
4	SG smoothing 17 points	6	0.3896	0.0431	0.0436	3.18E-08	1.28	0.6726	0.0285	0.0271	1.14E-02	2.06
5	SG smoothing 33 points	6	0.3894	0.0431	0.0436	3.97E-08	1.28	0.6712	0.0286	0.0271	1.15E-02	2.06
6	VN	5	0.3587	0.0441	0.0446	3.58E-08	1.25	0.6790	0.0283	0.0272	1.04E-02	2.05
7	MN	6	0.3815	0.0434	0.0438	3.97E-09	1.27	0.7120	0.0268	0.0253	1.09E-02	2.20
8	LBC + SNV	7	0.3908	0.0430	0.0435	3.97E-08	1.28	0.6564	0.0292	0.0272	1.28E-02	2.05
9	MSC+SG smoothing 3pts	5	0.3719	0.0437	0.0442	4.11E-08	1.26	0.6320	0.0303	0.0295	1.02E-02	1.89
10	MSC + 1st Der (3 points)	6	0.3884	0.0431	0.0436	4.37E-08	1.28	0.6554	0.0293	0.0268	1.37E-02	2.08
11	MSC + 2nd Der (3 points)	8	0.4730	0.0400	0.0405	4.50E-08	1.38	0.5937	0.0318	0.0317	8.43E-03	1.76
12	1st Der (3 points) + VN	6	0.3733	0.0436	0.0441	3.97E-08	1.26	0.6467	0.0297	0.0271	1.39E-02	2.05
13	2nd Der (3 points) + VN	8	0.3721	0.0437	0.0442	4.11E-08	1.26	0.5375	0.0339	0.0348	4.48E-03	1.60
14	1st Der (17 points) + VN	9	0.5810	0.0357	0.0361	4.11E-08	1.54	0.6024	0.0315	0.0295	1.34E-02	1.89

Appendix Table C10 Statistical results of fiber length NIR model in range 1000-2500 nm obtained by PLSR with different spectral pre-processing methods

No.	NIR spectral pre-processing	No. of PC	Calibration set (N=45)					Validation set (N=15)				
			R^2	RMSEC	SEC	Bias	RPD_c	Adjusted R^2	RMSEP	SEP	Bias	RPD_p
1	No pre-processing	5	0.4016	0.0426	0.0431	3.84E-08	1.29	0.6860	0.0280	0.0264	1.15E-02	2.11
2	MSC	4	0.4157	0.0421	0.0426	3.97E-08	1.31	0.6457	0.0297	0.0283	1.16E-02	1.97
3	SG smoothing 3 points	5	0.4018	0.0426	0.0431	3.31E-08	1.29	0.6853	0.0280	0.0264	1.15E-02	2.11
4	SG smoothing 17 points	5	0.4023	0.0426	0.0431	3.97E-08	1.29	0.6793	0.0283	0.0267	1.16E-02	2.09
5	SG smoothing 33 points	5	0.4015	0.0426	0.0431	3.18E-08	1.29	0.6756	0.0284	0.0268	1.16E-02	2.08
6	VN	5	0.4346	0.0414	0.0419	3.84E-08	1.33	0.7027	0.0272	0.0262	1.00E-02	2.13
7	MN	5	0.4316	0.0416	0.0420	-1.46E-08	1.33	0.7322	0.0258	0.0246	1.01E-02	2.27
8	LBC + SNV	3	0.3284	0.0452	0.0457	4.11E-08	1.22	0.6598	0.0291	0.0288	8.40E-03	1.93
9	MSC+SG smoothing 3pts	4	0.4156	0.0421	0.0426	5.83E-08	1.31	0.6455	0.0297	0.0283	1.16E-02	1.97
10	MSC + 1st Der (3 points)	2	0.5347	0.0376	0.0380	3.58E-08	1.47	0.6733	0.0285	0.0270	1.16E-02	2.07
11	MSC + 2nd Der (3 points)	3	0.8365	0.0223	0.0225	4.77E-08	2.47	0.6235	0.0306	0.0305	8.25E-03	1.83
12	1st Der (3 points) + VN	2	0.5479	0.0371	0.0375	4.50E-08	1.49	0.7139	0.0267	0.0250	1.13E-02	2.23
13	2nd Der (3 points) + VN	3	0.7498	0.0276	0.0279	3.18E-08	2.00	0.5424	0.0337	0.0325	1.25E-02	1.72
14	1st Der (17 points) + VN	2	0.4911	0.0393	0.0398	3.44E-08	1.40	0.6935	0.0276	0.0249	1.36E-02	2.24

Appendix Table C11 Statistical results of fiber length NIR model in range 1400-2500 nm obtained by PCR with different spectral pre-processing methods

No.	NIR spectral pre-processing	No. of PC	Calibration set (N=45)					Validation set (N=15)				
			R^2	RMSEC	SEC	Bias	RPD_c	Adjusted R^2	RMSEP	SEP	Bias	RPD_p
1	No pre-processing	5	0.3379	0.0449	0.0454	4.50E-08	1.23	0.5970	0.0317	0.0289	1.50E-02	1.93
2	MSC	4	0.3797	0.0434	0.0439	4.24E-08	1.27	0.6265	0.0305	0.0294	1.10E-02	1.89
3	SG smoothing 3 points	5	0.3380	0.0448	0.0454	3.97E-08	1.23	0.5974	0.0317	0.0288	1.50E-02	1.93
4	SG smoothing 17 points	5	0.3394	0.0448	0.0453	3.71E-09	1.23	0.6006	0.0315	0.0287	1.50E-02	1.94
5	SG smoothing 33 points	5	0.3405	0.0448	0.0453	3.97E-08	1.23	0.6013	0.0315	0.0286	1.51E-02	1.95
6	VN	4	0.3161	0.0456	0.0461	3.58E-08	1.21	0.5816	0.0323	0.0285	1.68E-02	1.95
7	MN	11	0.5161	0.0383	0.0388	2.68E-07	1.44	0.6769	0.0284	0.0269	1.13E-02	2.07
8	LBC + SNV	7	0.4059	0.0425	0.0430	4.50E-08	1.30	0.6270	0.0305	0.0283	1.35E-02	1.97
9	MSC+SG smoothing 3pts	4	0.3796	0.0434	0.0439	3.84E-08	1.27	0.6263	0.0305	0.0294	1.10E-02	1.89
10	MSC + 1st Der (3 points)	5	0.3877	0.0431	0.0436	3.71E-08	1.28	0.6045	0.0314	0.0275	1.66E-02	2.02
11	MSC + 2nd Der (3 points)	5	0.3762	0.0435	0.0440	3.71E-08	1.27	0.5857	0.0321	0.0311	1.12E-02	1.79
12	1st Der (3 points) + VN	6	0.5385	0.0374	0.0379	4.24E-08	1.47	0.5548	0.0333	0.0320	1.23E-02	1.74
13	2nd Der (3 points) + VN	9	0.3541	0.0443	0.0448	4.24E-08	1.24	0.5437	0.0337	0.0342	6.86E-03	1.63
14	1st Der (17 points) + VN	7	0.6199	0.0340	0.0344	3.44E-08	1.62	0.5740	0.0326	0.0311	1.25E-02	1.79

Appendix Table C12 Statistical results of fiber length NIR model in range 1400-2500 nm obtained by PLSR with different spectral pre-processing methods

No.	NIR spectral pre-processing	No. of PC	Calibration set (N=45)					Validation set (N=15)				
			R^2	RMSEC	SEC	Bias	RPD_c	Adjusted R^2	RMSEP	SEP	Bias	RPD_p
1	No pre-processing	5	0.4618	0.0404	0.0409	5.17E-08	1.36	0.6441	0.0298	0.0278	1.29E-02	2.01
2	MSC	4	0.4294	0.0416	0.0421	6.23E-08	1.32	0.6447	0.0297	0.0285	1.12E-02	1.96
3	SG smoothing 3 points	5	0.4598	0.0405	0.0410	2.12E-08	1.36	0.6507	0.0295	0.0275	1.28E-02	2.03
4	SG smoothing 17 points	5	0.4415	0.0412	0.0417	3.58E-08	1.34	0.6776	0.0283	0.0261	1.29E-02	2.14
5	SG smoothing 33 points	5	0.4263	0.0418	0.0422	2.65E-08	1.32	0.6818	0.0281	0.0257	1.32E-02	2.17
6	VN	4	0.4249	0.0418	0.0423	3.84E-09	1.32	0.6689	0.0287	0.0257	1.44E-02	2.17
7	MN	4	0.4159	0.0421	0.0426	-7.02E-08	1.31	0.6765	0.0284	0.0239	1.64E-02	2.33
8	LBC + SNV	2	0.3095	0.0458	0.0463	4.37E-08	1.20	0.6318	0.0303	0.0300	8.69E-03	1.86
9	MSC+SG smoothing 3pts	4	0.4294	0.0416	0.0421	6.09E-08	1.32	0.6445	0.0297	0.0285	1.12E-02	1.95
10	MSC + 1st Der (3 points)	1	0.3369	0.0449	0.0454	4.37E-08	1.23	0.5659	0.0329	0.0288	1.76E-02	1.94
11	MSC + 2nd Der (3 points)	5	0.9524	0.0120	0.0122	4.11E-08	4.59	0.5905	0.0319	0.0313	1.02E-02	1.78
12	1st Der (3 points) + VN	3	0.6696	0.0317	0.0320	4.11E-08	1.74	0.5755	0.0325	0.0307	1.32E-02	1.81
13	2nd Der (3 points) + VN	3	0.7140	0.0295	0.0298	4.11E-08	1.87	0.5183	0.0346	0.0352	6.55E-03	1.58
14	1st Der (17 points) + VN	3	0.5888	0.0353	0.0357	3.71E-08	1.56	0.6350	0.0301	0.0274	1.44E-02	2.03

Appendix Table C13 Statistical results of fiber coarseness NIR model in range 1000-2500 nm obtained by PCR with different spectral pre-processing methods

No.	NIR spectral pre-processing	No. of PC	Calibration set (N=45)					Validation set (N=15)				
			R^2	RMSEC	SEC	Bias	RPD_c	Adjusted R^2	RMSEP	SEP	Bias	RPD_p
1	No pre-processing	7	0.6402	0.0145	0.0147	5.13E-09	1.67	0.4271	0.0188	0.0194	9.35E-04	1.34
2	MSC	9	0.6198	0.0149	0.0151	2.09E-08	1.62	0.6248	0.0152	0.0157	3.06E-04	1.65
3	SG smoothing 3 points	7	0.6399	0.0145	0.0147	4.47E-09	1.67	0.4275	0.0188	0.0194	8.87E-04	1.34
4	SG smoothing 17 points	7	0.6369	0.0146	0.0147	-1.66E-10	1.66	0.4317	0.0187	0.0194	4.76E-04	1.34
5	SG smoothing 33 points	7	0.6343	0.0146	0.0148	4.64E-09	1.65	0.4336	0.0187	0.0193	1.74E-04	1.34
6	VN	11	0.6587	0.0141	0.0143	1.66E-09	1.71	0.6007	0.0157	0.0162	8.02E-04	1.60
7	MN	17	0.7511	0.0121	0.0122	-1.48E-07	2.00	0.7364	0.0127	0.0132	2.10E-04	1.97
8	LBC + SNV	9	0.6165	0.0150	0.0151	3.89E-09	1.61	0.6361	0.0150	0.0155	5.28E-04	1.68
9	MSC+SG smoothing 3pts	9	0.6195	0.0149	0.0151	1.83E-08	1.62	0.6246	0.0152	0.0157	3.02E-04	1.65
10	MSC + 1st Der (3 points)	6	0.4080	0.0186	0.0188	2.32E-09	1.30	0.3617	0.0198	0.0205	5.39E-04	1.27
11	MSC + 2nd Der (3 points)	4	0.2555	0.0208	0.0211	1.33E-09	1.16	0.4176	0.0189	0.0196	-6.50E-04	1.33
12	1st Der (3 points) + VN	5	0.4043	0.0186	0.0189	2.15E-09	1.30	0.4029	0.0192	0.0198	7.34E-04	1.31
13	2nd Der (3 points) + VN	5	0.2702	0.0206	0.0209	1.16E-09	1.17	0.4236	0.0188	0.0195	-8.63E-04	1.33
14	1st Der (17 points) + VN	7	0.4946	0.0172	0.0174	6.62E-10	1.41	0.4575	0.0183	0.0189	1.76E-04	1.37

Appendix Table C14 Statistical results of fiber coarseness NIR model in range 1000-2500 nm obtained by PLSR with different spectral pre-processing methods

No.	NIR spectral pre-processing	No. of PC	Calibration set (N=45)					Validation set (N=15)				
			R^2	RMSEC	SEC	Bias	RPD_c	Adjusted R^2	RMSEP	SEP	Bias	RPD_p
1	No pre-processing	4	0.5902	0.0155	0.0156	4.47E-09	1.56	0.4854	0.0178	0.0184	-5.92E-04	1.41
2	MSC	7	0.6649	0.0140	0.0141	2.52E-08	1.73	0.6510	0.0147	0.0152	1.14E-04	1.71
3	SG smoothing 3 points	4	0.5878	0.0155	0.0157	3.81E-09	1.56	0.4824	0.0179	0.0185	-6.72E-04	1.41
4	SG smoothing 17 points	4	0.5749	0.0157	0.0159	1.66E-09	1.53	0.4628	0.0182	0.0188	-1.18E-03	1.38
5	SG smoothing 33 points	5	0.6460	0.0144	0.0145	4.47E-09	1.68	0.4621	0.0182	0.0188	5.61E-04	1.38
6	VN	7	0.6344	0.0146	0.0148	2.05E-08	1.65	0.6232	0.0152	0.0156	2.24E-03	1.67
7	MN	8	0.6822	0.0136	0.0138	2.34E-07	1.77	0.6888	0.0138	0.0143	8.07E-04	1.82
8	LBC + SNV	8	0.6977	0.0133	0.0134	-7.12E-09	1.82	0.6190	0.0153	0.0158	8.86E-04	1.64
9	MSC+SG smoothing 3pts	7	0.6627	0.0140	0.0142	1.41E-08	1.72	0.6503	0.0147	0.0152	3.43E-05	1.71
10	MSC + 1st Der (3 points)	2	0.5600	0.0160	0.0162	2.15E-09	1.51	0.4160	0.0190	0.0196	1.34E-03	1.33
11	MSC + 2nd Der (3 points)	1	0.4399	0.0181	0.0183	2.48E-09	1.34	0.4584	0.0183	0.0188	1.74E-03	1.38
12	1st Der (3 points) + VN	2	0.5576	0.0161	0.0162	2.15E-09	1.50	0.4378	0.0186	0.0192	1.26E-03	1.35
13	2nd Der (3 points) + VN	1	0.4380	0.0181	0.0183	2.48E-09	1.33	0.4269	0.0188	0.0194	-3.36E-04	1.34
14	1st Der (17 points) + VN	3	0.6209	0.0149	0.0150	2.48E-09	1.62	0.4956	0.0176	0.0182	-5.28E-04	1.43

Appendix Table C15 Statistical results of fiber coarseness NIR model in range 1400-2500 nm obtained by PCR with different spectral pre-processing methods

No.	NIR spectral pre-processing	No. of PC	Calibration set (N=45)					Validation set (N=15)				
			R^2	RMSEC	SEC	Bias	RPD_c	Adjusted R^2	RMSEP	SEP	Bias	RPD_p
1	No pre-processing	11	0.6364	0.0146	0.0147	-3.81E-09	1.66	0.5843	0.0160	0.0165	-1.31E-03	1.57
2	MSC	11	0.6209	0.0149	0.0150	1.33E-08	1.62	0.6030	0.0156	0.0162	4.41E-04	1.61
3	SG smoothing 3 points	11	0.6364	0.0146	0.0147	1.31E-08	1.66	0.5848	0.0160	0.0165	-1.32E-03	1.58
4	SG smoothing 17 points	11	0.6368	0.0146	0.0147	-3.64E-08	1.66	0.5905	0.0159	0.0164	-1.50E-03	1.59
5	SG smoothing 33 points	11	0.6441	0.0144	0.0146	7.29E-09	1.68	0.6005	0.0157	0.0161	-1.70E-03	1.61
6	VN	9	0.6859	0.0135	0.0137	1.99E-09	1.78	0.5049	0.0175	0.0181	-3.46E-04	1.44
7	MN	16	0.7199	0.0128	0.0129	-2.78E-07	1.89	0.6113	0.0155	0.0160	-5.47E-04	1.62
8	LBC + SNV	9	0.6151	0.0150	0.0152	4.31E-09	1.61	0.6384	0.0149	0.0155	1.32E-04	1.68
9	MSC+SG smoothing 3pts	11	0.6206	0.0149	0.0150	1.71E-08	1.62	0.6025	0.0157	0.0162	4.08E-04	1.61
10	MSC + 1st Der (3 points)	8	0.5410	0.0164	0.0165	6.62E-10	1.48	0.4389	0.0186	0.0192	1.40E-03	1.35
11	MSC + 2nd Der (3 points)	5	0.3992	0.0187	0.0189	2.82E-09	1.29	0.3838	0.0195	0.0202	3.69E-04	1.29
12	1st Der (3 points) + VN	6	0.5582	0.0161	0.0162	9.93E-10	1.50	0.4156	0.0190	0.0196	5.79E-04	1.32
13	2nd Der (3 points) + VN	5	0.3515	0.0195	0.0197	2.32E-09	1.24	0.4270	0.0188	0.0194	-4.51E-04	1.34
14	1st Der (17 points) + VN	8	0.5565	0.0161	0.0163	1.66E-10	1.50	0.3989	0.0192	0.0199	1.16E-03	1.31

Appendix Table C16 Statistical results of fiber coarseness NIR model in range 1400-2500 nm obtained by PLSR with different spectral pre-processing methods

No.	NIR spectral pre-processing	No. of PC	Calibration set (N=45)					Validation set (N=15)				
			R^2	RMSEC	SEC	Bias	RPD_c	Adjusted R^2	RMSEP	SEP	Bias	RPD_p
1	No pre-processing	9	0.7371	0.0124	0.0125	-3.31E-08	1.95	0.6321	0.0151	0.0156	3.28E-04	1.67
2	MSC	7	0.6409	0.0145	0.0146	-5.05E-09	1.67	0.6269	0.0152	0.0157	-3.02E-04	1.66
3	SG smoothing 3 points	9	0.7372	0.0124	0.0125	-6.95E-09	1.95	0.6305	0.0151	0.0156	3.32E-04	1.67
4	SG smoothing 17 points	8	0.7091	0.0130	0.0132	-4.17E-08	1.85	0.6360	0.0150	0.0155	-3.03E-04	1.68
5	SG smoothing 33 points	9	0.7051	0.0131	0.0133	1.61E-08	1.84	0.6339	0.0150	0.0155	-6.84E-04	1.67
6	VN	8	0.7366	0.0124	0.0125	-1.11E-08	1.95	0.5301	0.0170	0.0176	2.29E-04	1.48
7	MN	9	0.7927	0.0110	0.0111	-1.96E-07	2.20	0.6125	0.0155	0.0160	-1.10E-03	1.63
8	LBC + SNV	7	0.6785	0.0137	0.0138	-6.62E-09	1.76	0.6249	0.0152	0.0157	5.45E-04	1.65
9	MSC+SG smoothing 3pts	7	0.6408	0.0145	0.0146	-2.98E-09	1.67	0.6278	0.0151	0.0157	-3.26E-04	1.66
10	MSC + 1st Der (3 points)	7	0.9248	0.0066	0.0067	8.28E-10	3.65	0.5268	0.0171	0.0176	-9.72E-04	1.47
11	MSC + 2nd Der (3 points)	2	0.7114	0.0130	0.0131	3.81E-09	1.86	0.3923	0.0193	0.0196	3.88E-03	1.33
12	1st Der (3 points) + VN	3	0.6537	0.0142	0.0144	1.16E-09	1.70	0.4330	0.0187	0.0193	-2.96E-04	1.34
13	2nd Der (3 points) + VN	2	0.5918	0.0154	0.0156	2.65E-09	1.57	0.4400	0.0186	0.0192	1.08E-03	1.35
14	1st Der (17 points) + VN	3	0.5693	0.0159	0.0160	4.97E-10	1.52	0.4248	0.0188	0.0194	-1.19E-03	1.34

CIRRICULUM VITAE

NAME	: Mr. Worapong Athihirunwong
BIRTH DATE	: April 17, 1984
BIRTH PLACE	: Bangkok, Thailand
EDUCATION	: 2007 - B.Sc. (Pulp and Paper Technology) Kasetsart University
SCHOLARSHIP/AWARDS	: 2008 - Assistant Teaching Scholarship, The Graduate School Kasetsart University. 2010 - The Capacity Building for KU Students on Internationalization project, Kasetsart University 2010 - The Best Speaker Award, The 2nd Graduate Student Seminar 2010, Faculty of Forestry, Kasetsart University

.