INVESTIGATION OF NIR SPECTROSCOPY FOR IDENTIFYING AND SORTING WOOD WITH RESPECT TO SPECIES, MOISTURE CONTENT, AND WEATHERING

by

Ekaterina Tounis
B. Applied Chemistry and Biology, Ryerson University, 2007

A Thesis

presented to Ryerson University

in partial fulfillment of the

requirements for the degree of

Master of Science

in the Program of

Molecular Science

Toronto, Ontario, Canada, 2009

I hereby declare that I am the sole author of this thesis.

I authorize Ryerson University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

____ _____

I further authorize Ryerson University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

_____ _____

Abstract

INVESTIGATION OF NIR SPECTROSCOPY FOR IDENTIFYING AND SORTING WOOD WITH RESPECT TO SPECIES, MOISTURE CONTENT, AND WEATHERING

by

Ekaterina Tounis

An abstract submitted to Ryerson University in partial fulfillment of the requirements for the degree of Master of Science in the Program of Molecular Science, 2009.

Near-infrared spectroscopy can characterize wood surfaces fast and without significant surface preparation. It is based on molecular overtone and combination vibrations which are typically very broad, leading to complex spectra. Multivariate calibration techniques are often employed to extract the desired chemical information. This study focused on the potential of near-infrared spectroscopy combined with partial least squares for identifying and sorting wood with respect to species and physical properties and on the effects of wood preparation and weathering on the precision of analysis. It was shown that a range of moisture content values and artificial weathering periods could be well predicted independently of wood species analyzed. Species within the spruce-pine-fir species group could be predicted reasonably well when tested under similar conditions. When different moisture contents and weathering exposure periods were introduced, species prediction was still possible, but, with decreased prediction ability.

Acknowledgments

I had the privilege to be supervised by great people, who offered me opportunity and help in my studies. I am really grateful to Professors Darrick Heyd, Paul Cooper, and Stephen Wylie whose assistance, advice and supervision were very useful. I also appreciate advice and information provided by Brigitte Leblon from UNB. And I gratefully acknowledge the assistance of Dragica Jeremic and Tony Ung for contributing their time and expertise.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

CCA	Chromated Copper Arsenate
DS	Degree of Substitution
EM	Electromagnetic
EMC	Equilibrium Moisture Content
E_L	Longitudinal modulus of Elasticity
Fig	Figure
FSP	Fiber Saturation Point
FT	Fourier Transform
InGaAs	Indium Gallium Arsenide
IR	Infrared
KNN	K Nearest Neighbours
MC	Moisture Content
MFA	Microfibril Angle
MIR	Mid-Infrared
MLR	Multiple Linear Regression
MOE	Modulus of Elasticity
MOR	Modulus of Rupture
MSC	Multiplicative Scatter Correction
NIR	Near Infrared
NIRS	Near Infrared Spectroscopy
nm	nanometres
OSC	Orthogonal Signal Correction
PC	Principle Component
PCA	Principle Component Analysis
PCR	Principle Component Regression
PLS	Partial Least Square
PLS-DA	Partial Least Squares Discriminant Analysis
PLSR	Partial Least Square Regression
RH	Relative Humidity
RL	Radial-Longitudinal
RMSEC	Root Mean Square of Calibration
RMSEP	Root Mean Square Error of Prediction
RT	Radial-Transverse
S/G	Syringyl/Guaiacyl ratio
SEP	Standard Error of Predictions
SIMCA	Soft Independent Modeling of Class Analogy

SNV	Standard Normal Variate
SPF	Spruce-Pine-Fir
SS	SilviScan-2
UTS	Ultimate Tensile Stress
UV	Ultraviolet
Vis	Visible

1. Introduction

For many value-added applications, a rapid and reliable sorting of solid wood by species and other properties is needed. Near Infrared Spectroscopy (NIRS) has developed as a powerful tool to rapidly predict chemical and physical properties on the production line or through handheld instruments based on the surface chemistry of wood. One application of industrial interest is development of a fast and reliable sorting technology for separating fir, spruce, and pine. These three species are currently marketed as one (SPF) but it will be beneficial to separate them. Fir is more permeable than spruce and pine, and so longer drying times are needed and it should be separated from the other two species for drying and preservative treatment. Fir is also weaker than spruce or pine, and so is less suitable for construction. It is also beneficial to separate spruce and pine because spruce is selected for machine stress rating for lumber construction, while pine is better for wood treating or window frames and furniture.

This study focused on the potential of NIRS for separating spruce, fir, and pine, and on the effects of wood preparation and weathering of wood surfaces on the precision of analysis. Results suggest that NIRS has potential as a cost-effective and rapid method for identifying or separating wood species and wood of different properties.

1.1. Near Infrared Spectroscopy

1.1.1. Principles of NIR Spectroscopy

Near infrared (NIR) refers to the spectral region adjacent to the long-wavelength side of the visible (Vis) range. The NIR wavelength range lies between 750 and 2600 nm. The transitions involved in NIRS are usually the weak vibrational overtone and combination bands.

Molecules are capable of storing energy as vibrations; i.e., internal motion in which the nuclei oscillate about their equilibrium bond positions. In general, a non-linear molecule will have $3N$-6 normal modes, where N is the number of nuclei in the molecule. Many of the modes will be *infrared active*; i.e., they will be capable of absorbing infrared (IR) radiation, which process results in a change of the vibrational energy of the molecule. Normal vibrational modes that result in a change of dipole moment of the molecule are IR active; fully symmetric modes, which do not change the dipole moment, are not.

Each IR-active normal vibration of a molecule will absorb radiation at its fundamental frequency, v_i. The energy of the photon, hv_i, where h is Planck's constant, corresponds to the difference in energy between the lowest (or ground) vibrational level (n = 0, where n represents the vibrational quantum number and 0 denotes the lowest vibrational level) and the first excited vibrational state (n =1) of the mode; thus, when the photon is absorbed the molecule gains vibrational energy, as well as a series of overtones and

combination bands at higher energy. The fundamental transition is allowed by quantum mechanical rules (provided symmetry conditions are met) and is the strongest transition for the mode because the frequency of the oscillating dipole moment of the molecule matches that of the photon's electric field, allowing excellent transfer of energy. Fundamental frequencies are usually observed in the mid infrared spectral region (MIR; 2.5 to 15 μm).

Absorption of the more energetic NIR radiation by organic molecules is due to overtone and combination bands, primarily of O-H, C-H, N-H, C-C and C=O groups. The overtones involve transitions from the ground vibrational level to higher excited states: the first overtone is the transition from $n = 0$ to $n = 2$; the second overtone is $n = 0$ to $n = 3$, etc. Because they involve transitions to higher states, they are observed at higher energy in the spectrum, i.e. the NIR region. The overtones are formally forbidden by quantum mechanical selection rules because the frequency of the photon no longer matches that of the oscillating dipole moment (which is roughly half that of the photon), and so are much weaker than the fundamentals. Combination bands involve simultaneous excitation of two fundamentals and so they also appear in the NIR. They are also weak transitions, like the overtones, for a similar reason.

The NIR spectrum of an organic material includes overtone and combination bands from all the organic functional groups in the material. Because the structures of organic materials like wood are very complex, the resulting spectra can be difficult to interpret. When performing NIR analysis specific features in the spectra must be identified and

characterized by means of statistical methods. Chemometrics procedures are often necessary for both qualitative and quantitative analysis. Fortunately, several commercial software packages are available to accomplish that. For more about NIR principles see reference (1, 2).

1.1.2. Theory of NIR Absorption Spectroscopy

To understand the types of measurements possible using NIR, it is useful to understand several general properties of electromagnetic (EM) radiation, as well as the basics of classical molecular and atomic structure. EM radiation has wave-like properties. Frequency refers to the number of complete "waves" that occur in unit time (usually one second). Wavelength is the distance between equivalent points on successive waves. These parameters are related by:

$$c = v\lambda \qquad [1]$$

where c = speed of light in vacuum, λ = wavelength, and v = frequency in cycles per second (Hz). Light, as well as being a wave, consists of photons which have properties of both waves and particles. The energy of a photon is given by:

$$E=hv = hc/\lambda \qquad [2]$$

where E = energy and h = Planck's constant, so that the energy of a photon is related to the wavelength at which the radiation is emitted. The general properties of waves of certain energies allow their classification across the full EM spectrum.

Photon energy can be absorbed by molecules, elevating the energy of the molecule by a discrete amount to a higher energy level. The energy levels of a diatomic molecule can be roughly calculated using equation [3] which is based on a quantum mechanical harmonic oscillator.

$$E_n = \left(n + \frac{1}{2}\right)\frac{h}{2\pi}\sqrt{\frac{k}{\mu}} \qquad [3]$$

where E_n = the molecule vibrational energy, n = quantum number (n = 0,1,2,3 ...), h = Plank's constant, k = the force constant (which reflects bond strength), and μ = the reduced mass $\left(\frac{m_1 m_2}{m_1 + m_2}\right)$. A transition from n = 0 to n =1 is known as a fundamental absorption. The overtones are the result of forbidden transitions arising from the ground vibrational energy level to levels where $n > 1$.

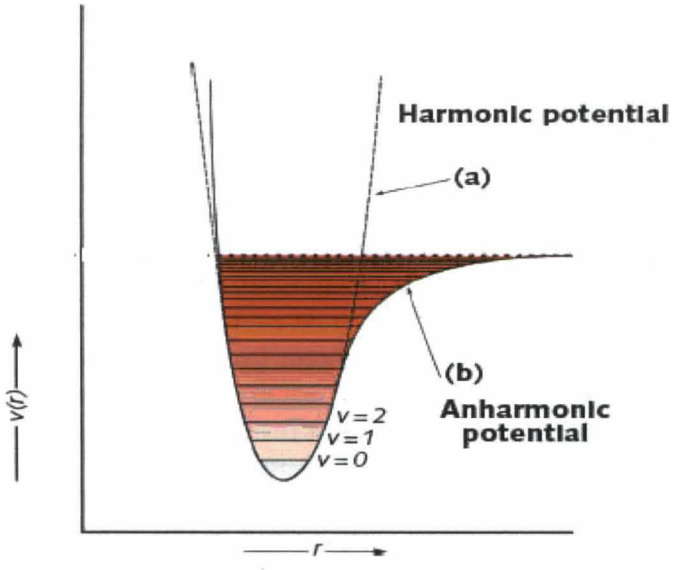
The majority of overtone bands in the NIR arise from X-H stretching modes. Due to the large mass difference between the two atoms, large amplitude vibrations arise with high anharmonicity and large dipole moments. The large anharmonicity causes the frequencies of the first and second overtones being slightly different from (usually less than) whole-number multiples of the fundamental frequency. As a result, first, second, and higher overtones usually occur at longer wavelengths than expected. This is mathematically described by equation [4]. This equation relates the observed frequency ν_n of the n^{th} harmonic to the frequency ν_o of the fundamental (1):

$$\nu_n = n\nu_o\,(1 - nx) \qquad [4]$$

where n is the vibrational quantum number (=1 for fundamental and =2 for second overtone) and x is a positive number indicative of the anharmonicity that measures the deviation of the potential function from the ideal parabolic function (**Fig.1.1**).

Unlike the harmonic oscillator, the energy levels expressed in equation [3] are no longer equidistant from each other (**Fig. 1.1**) and the selection rule is expanded to transitions over more that one energy level. Reference (1) gives the approximate areas where overtones of some of the most common absorbers may be expected to occur.

Fig. 1.1. The energy diagram of a molecule's vibrational model showing an (a) ideal diatomic or (b) anharmonic diatomic oscillator. (3)



The sample can be made of different types of molecules each with its own absorption spectrum, resulting in a complex spectrum. When a measurement is performed on this sample, the instrument is measuring the number of photons which undergo the absorption

process for a particular wavelength. The number of photons absorbed is proportional to the amount of a particular type of molecule present in the sample. This statement is Beer's law (equation [5]). In principle, this is the basis for an NIR measurement.

$$\log_{10} (1/T_{\lambda}) = \varepsilon_{\lambda} bc \qquad [5]$$

where ε_{λ} is the molar absorption coefficient that is a characteristic of each molecular species, b is the path length of the irradiating energy through the sample, and c is the analyte concentration. The quantity log $(1/ T_{\lambda})$ is 'absorbance' (A). In the field of NIRS, reflectance (R) is analogous to transmittance (T); equation [5] can thus be expressed in terms of reflectance as (1):

$$\log_{10} (1/R_{\lambda}) = \varepsilon_{\lambda} bc = A \qquad [6]$$

In NIRS the reflectance of solid samples is measured where the path length cannot be kept constant and will vary with sample. The net consequence of this scattering is to change the proportion of absorbed and reflected radiation so that path length becomes an extra unknown along with concentration. Therefore, even if only one component was known to absorb at just one wavelength, it would be necessary to have another measurement wavelength to solve a pair of simultaneous equations relating path length and concentration to absorbance at the two wavelengths. Features that affect path length include density, moisture content (MC) and absorption itself. In practice, most measurements require information from more than one region of the NIR spectrum, so that a matrix solution to a set of simultaneous equations forms part of the calibration process. Calibration is a regression modeling procedure that identifies the minimum subset of wavelength terms that best explains the chemical property across a population

of similar commodities showing multivariate changes in composition. For more about NIR theory see reference (1, 2).

1.1.3. NIR Spectra

All molecules have normal vibrations at characteristic wavelengths. In each normal mode of vibration, all of the functional groups of atoms of the same type in a molecule vibrate at a certain frequency range. There are two main types of molecular vibrations: stretching and bending. Stretching is movement along the axes of bonds, so that the distance between atoms changes rhythmically. Stretching vibrations are either symmetrical or asymmetrical. The absorption frequency of asymmetric stretch is higher than that of symmetric stretch. Bending vibrations may involve changes in bond angles between atoms, or movements of groups of atoms with respect to the rest of the molecule with minimal movement of the atoms in the group relative to one another. Bending vibrations are either out-of-plane or in-plane. Out-of-plane bending consists of wagging and twisting, and in-plane bending consists of scissoring and rocking. For any given molecule, stretching vibrations occur at the highest frequency, followed by scissoring, wagging, and twisting and rocking. Only vibrations that result in rhythmic changes in the dipole moment of a molecule (permanent displacement of electric charge) are capable of absorbing in the IR region.

Polyatomic molecules composed of N atoms have $3N$ degrees of freedom. For non-linear molecules, three of these describe translation of the molecule and another three

describe rotation. The remaining (3*N*-6) degrees of freedom describe vibrational motions, all of which are theoretically capable of absorbing photon energy, and so there are 3*N*-6 possible fundamental absorbencies for a non-linear molecule. For linear molecules, only two degrees of freedom are required to describe rotation because rotation about the molecular axis is not measurable, so that there are (3*N*-5) degrees of freedom available for vibrational motions. Of these, *N*-1 are stretching motions and 2*N*-5 are bending motions. There is more deformation or bending than stretching vibrations, but stretching vibrations are usually the more intense.

The above rules for the number of absorption bands apply only to the number of fundamental bands. It does not apply to the number of overtone and combination bands. Combination bands occur when the absorbed photon excites two or more vibrations simultaneously. For this to happen, the energy of the photon has to be equal the sum of the energies of the coupling vibrations. They are also influenced by the anharmonicity of the vibration which increases with the amplitude of atomic oscillation. In practice, the number of fundamental bands observed is rarely exactly 3*N*-6 for any molecule. The number can be decreased by several factors, including a high degree of molecular symmetry. Some fundamentals occur so close together that they are not separated by the spectrometer, and others are too weak to be observed. Finally, some bands occur outside of the range of the spectrometer.

If the sample being irradiated was totally reflective and contained molecules all of which were 100% nonpolar and symmetrical, the spectrum would theoretically appear as

a straight line across the wavelength range. In practice, at each wavelength point there are molecules that are capable of absorbing energy. Each individual constituent present in any sample, will absorb energy according to Beer's law to a greater or lesser degree at different wavelength points, depending upon their concentration. Accordingly, the amount of absorbance/ reflectance that takes place will differ at different wavelength points, thereby creating the undulating pattern as the spectrum. The shape of the spectral trace in terms of "peaks and valleys" of absorbance is characteristic of all of the absorbing molecules present in the sample and is affected by the composition of the molecules, by the presence and magnitude of dipoles, by interactions between molecules, by the resolution capability of the spectrometer, by the efficiency of the detector at different wavelengths, and by other factors.

Absorption of NIR radiation can be attributed largely to the low mass of the hydrogen atom. Not only does this low mass cause resonant molecular vibration at high frequency, but the resultant oscillations are often anharmonic; and since anharmonicity in a mode enhances overtone excitation these high-frequency vibrations exhibit strong overtone absorption at near multiples of the fundamental frequency. Most of the fundamental oscillations of the bound hydrogen atom fall in the wavelength range of 2700-3600 nm (IR region). Most overtones are found in the area from 400 to 2200 nm (Vis and NIR region). The area between 1800 and 2500 nm is the combination area. All the information about NIR spectra is described in reference (1, 2).

1.1.4. Chemometric Models

The step in NIR analysis entailing the most planning preparation is the collection of the calibration samples, termed the training set. A very important step in achieving success is ensuring the samples have been analyzed as accurately and precisely as conventional techniques allow. These analyses are termed reference analyses. In order to make quantitative measurements or qualitative discriminations, one must have a computer with access to one or more chemometric model.

A chemometric model is developed by measuring a group of samples that display the maximum variability of the characteristic of interest by NIRS. The same samples are also used to measure the characteristic of interest by a reference method. The spectral data and independent test data are then analyzed using chemometrics software. An example of a statistical process used in quantitative or qualitative spectral analysis is partial least squares (PLS). When a sufficient number of samples have been collected and analyzed, a mathematical model is constructed that describes the relationship between certain spectral features and the sample characteristic of interest. Subsequently, that same characteristic in a new target samples can be measured by applying the chemometric model to the spectrum of the target samples.

Because there is a multitude of variables (for example moisture content, sapwood-heartwood differences, density within species, and perhaps surface roughness) that may all confound the signals, analytical refinements may be necessary to optimize the

analysis. Examples are first and second order derivatives and multiplicative scatter correction (MSC) transformation. MSC can be used on spectroscopic data when calculating PLS models to reduce their complexity as it separates the multiplicative or scattering variations from the chemical information. Scattering phenomena can arise from variations in, for example, water content and surface roughness in the samples, which are of relevance to this study. (2, 4)

1.1.5. NIR Instrumentation

The discovery of NIR energy is ascribed to Herschel in 1800, but the first industrial application began in the 1950s. Classical spectroscopists have traditionally avoided the NIR: Researchers were convinced that overtones and combination absorptions that occurred in the NIR were of little consequence because of their weak absorption cross-sections and broad overlapping resonances. The development of sensitive photon detectors revived interest in the NIR, and the technology was attached to UV-Vis instruments. This continued the separation of NIR from the mainstream of IR research. However, in the 1960s, the advent of the minicomputer enabled researchers to look more closely at the Vis and NIR spectra of various products for the purpose of sorting according to quality and composition. In the 1980s, a single unit stand-alone NIRS system became available, and the use of NIRS was focused mainly on chemical analysis. With the developments of light-fiber optics in the mid 80s and the monochromator-detector in early 90s, NIRS became a more powerful technique for scientific research, for example, for research involving remote measurements.

Instrumentation for NIRS is similar to traditional, scanning UV-Vis spectrophotometers. It has a radiation source, a detector, and a dispersive element. Incandescent (emission of light from a hot body) or quartz halogen light bulbs are most often used as broadband sources of NIR radiation. Light-emitting diodes which offer greater lifetime and spectral stability are also used. Wavelength isolators range in technology and include diode arrays, laser arrays, filters, and monochromators (prisms or gratings).

Fourier transform (FT)-NIR instruments using an interferometer are also available. They fall into one of two groups: 1) discrete-value and 2) full-spectrum devices. Full spectrum spectrometers produce spectra with equally spaced data across the full range. Discrete-value spectrometers are described as "discrete" because discrete wavelengths are used rather than a continuum of light.

The type of detector employed depends mainly on the range of wavelengths to be measured. Silicon-based charge-coupled detectors are suitable for the shorter end of the NIR range. Indium gallium arsenide (InGaAs) and lead (II) sulfide (PbS) detectors are sufficiently sensitive over most of the NIR range. In some diode array NIRS instruments, both silicon-based and InGaAs detectors are used in the same instrument. Such instruments can measure both Vis and NIR spectra simultaneously. For more about NIR instrumentation see reference (1).

1.1.6. Advantages and Disadvantages of NIRS

The biggest advantage of NIR is that it is suitable for in-situ analysis with little or no sample preparation, and it is rapid enough for near-real time analysis. Unlike most conventional analytical methods such as chromatography, NIRS is rapid (a spectrum can be acquired in as little as a tenth of a second), non-destructive, does not require chemicals, or generate chemical wastes requiring disposal, simultaneously determines numerous constituents or parameters, and can be transported to nearly any environment as it is truly portable for field work. NIR instrumentation is simple to operate by non-chemists, and operates without fume hoods, drains, or other installations.

Because the absorbance in the NIR region is lower than in neighbouring regions and generally obeys Beer's law, it is possible to analyze bulk samples without the need for dilution or other elaborate sample preparation. Thus, the results provided by NIR are typically more representative than those provided by other analytical means.

NIR is not useful as a stand-alone technology for quantitative analysis. Separate determinations by an independent and reliable method are required for each constituent or parameter to be determined with the training set and model, and a portion of unknown samples must be analyzed periodically, often by expensive and complicated reference methods, to ensure that calibrations remain reliable. The calibration methods rely on sophisticated chemometric techniques, thus requiring personnel who are trained in chemometrics. The above information was taken from reference (2).

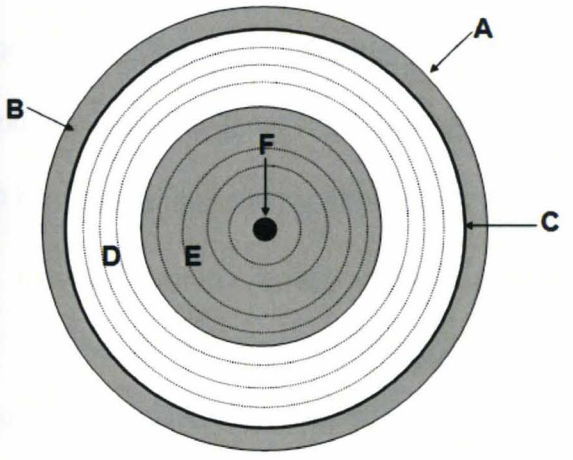
1.2. Wood Structure and Chemistry

1.2.1. Macro-structure of Wood

Wood is fibrous and lignified structural tissue produced as secondary xylem in the stems of woody plants. In a living tree, it conducts water and nutrients to the leaves and other growing tissues, and has a support function, enabling plants to reach large sizes. A cross section of a tree (**Fig. 1.2**) shows the following features: bark, which is divided into an outer corky dead part (A), whose thickness varies with species and age of trees, and an inner thin living part called phloem (B), which carries food from the leaves to growing parts of the tree; wood or xylem, which is differentiated into sapwood (D) and heartwood (E). Sapwood contains both living and dead tissue and carries sap from the roots to the leaves. Heartwood is formed by a gradual change in the sapwood and is inactive; and pith (F), a small core of tissue located at the center of tree stems, branches, and twigs about which initial wood growth takes place. The cambium layer (C), which is inside the inner bark forms wood and bark cells. New wood is laid down to the outside of old wood and the diameter of the woody trunk increases. Sapwood which is located between the cambium and heartwood (**Fig. 1.2**D) may vary in thickness and number of growth rings depending on the wood species and growth conditions. More vigorously growing trees have wider sapwood. The transition from sapwood to heartwood is accompanied by an increase in extractive content. The basic strength of the wood is essentially not affected by this transition.

In most species in temperate climates, the difference between wood that is formed early in a growing season and that formed later is sufficient to produce well-marked annual growth rings. The age of a tree may be determined by counting these rings. Nevertheless, if the growth in diameter is interrupted, more than one ring may be formed in the same season and trees that have only very small crowns or that have accidentally lost most of their foliage may form an incomplete growth layer. The inner part of the growth ring formed first in the growing season is called earlywood and the outer part formed later in the growing season, is called latewood. Actual time of formation of these parts of a ring may vary with environmental and weather conditions. Earlywood is characterized by cells with relatively large cavities and thin walls. Latewood cells have smaller cavities and thicker walls. When growth rings are prominent, earlywood has different physical properties than latewood. Earlywood is lighter in weight, softer, and weaker than latewood. Because of the greater density of latewood, the proportion of latewood is sometimes used to evaluate the strength of the wood. For in depth description of macro-structure of wood sees reference (5, 6).

Fig. 1.2. Cross section of tree trunk: (A) outer bark, (B) Phloem, (C) cambium, (D) sapwood, (E) heartwood, and (F) pith



1.2.2. Micro-structure of Wood

Trees are classified into two groups: hardwoods (angiosperms or flowering plants) and softwoods (gymnosperms or conifers or needle-bearing trees). The difference is in the cellular structure of the wood. The microscopic structure of wood generally resembles a bundle of straws glued together; each straw represents a cell with a lingo-cellulosic wall and a hollow centre (lumen), and lots of fine holes through the walls. Wood cells are of various sizes and shapes and are tightly bound together. The majority (over 90%) of wood cells are longitudinally oriented (parallel to the tree trunk); these cells are called fibers or tracheids. Their length is highly variable within a tree and among species. Hardwood fibers average about 1000 μm in length; softwood fibers range from 2000 to 5000 μm in length, 50 to 60 μm in width and have a cell wall thickness of two to eight μm. Wood fibers provide strength. In addition to fibers, hardwoods have cells of relatively large diameter known as vessels or pores. These cells function as pipes that

move sap up the tree through the mass of fibres. With some experience it is possible to distinguish different hardwoods based on the number, size and location of the vessels. Softwoods do not contain vessels for conducting sap longitudinally in the tree; this is done by the tracheids. Both hardwoods and softwoods have cells that are oriented horizontally in the direction from pith toward bark. These cells conduct sap radially across the grain and are called ray parenchyma cells. Another type of wood cell, known as longitudinal or axial parenchyma cells, functions mostly in the storage of food.

Dry wood is mainly composed of cellulose, lignin, hemicelluloses, and minor amounts (usually 5% to 10%) of extraneous materials. Cellulose, the major component, constitutes about 50% of wood by weight. It is a high-molecular-weight linear polymer consisting of chains of several hundred to over ten thousand β-linked glucose monomers. During growth of the tree, the cellulose molecules are arranged into ordered strands called fibrils, which in turn are organized into the larger structural elements that constitute the cell wall of wood cells. Most of the cell wall cellulose is crystalline.

Lignin constitutes 23% to 33% of the wood in softwoods and 16% to 25% in hardwoods. Although lignin occurs in wood throughout the cell wall, it is concentrated toward the outside of the cells and between cells. Lignin binds individual cells together. Lignin is a three-dimensional phenylpropanol polymer. In general, the lignins of softwoods are made up entirely of guaiacyl nuclei, whilst the lignins of hardwoods consist of a mixture of guaiacyl and syringyl nuclei.

The hemicelluloses are associated with cellulose and are branched, low-molecular-weight polymers composed of several different kinds of pentose and hexose sugar monomers. The relative amounts of these sugars vary with species. The total cellulose and hemicellulose content is known as holocellulose, and is generally found to be higher in hardwoods than softwoods.

Unlike the major constituents of wood, extraneous materials are not structural components. Both organic and inorganic extraneous materials are found in wood. The organic component is called the extractives, which contribute to such wood properties as color, odour, taste, decay resistance, density, hygroscopicity, and flammability. Extractives include tannins and other polyphenolics, pigments, essential oils, fats, resins, waxes, starch, and simple metabolic intermediates. The component is called extractives because it can be removed from wood by extraction with hot water and organic solvents. Extractives may constitute approximately 5% to 30% of the wood, depending on factors such as species, growth conditions, and time of year when the tree is cut. The inorganic component of extraneous material (ash) generally constitutes 0.2% to 1.0% of the wood. Calcium, potassium, and magnesium are the more abundant elemental constituents. Trace amounts of phosphorus, sodium, iron, silicon, manganese, copper, zinc, and a few other elements are generally present. For in depth description of micro-structure of wood see reference (5, 6).

1.3. Literature Review- Application of NIRS to Wood

1.3.1. Chemical Composition of Wood

Wood is a composite material consisting of three major polymers, cellulose, hemicellulose, and lignin, which serve as skeletal, matrix, and encrusting substances, respectively. The extractives are usually low molecular weight, extracellular compounds. Considerable NIR research concerning these compounds has been published in the area of non-destructive evaluation with the aid of multivariate analysis.

The chemical components of wood entering the kraft pulping process are of high interest for many reasons. Traditional methods which rely on wet chemistry are slow and expensive, restricting the number of samples that may be processed. NIRS is a promising method that can be adapted for rapid measurements on wood. However, the reference measurements must be accurate and must represent a wide range of values to achieve valid predictions. (8)

1.3.1.1. Hardwoods

Raymond and Schimleck (2002) examined the feasibility and efficiency of predicting cellulose content using NIR reflectance analysis for *Eucalyptus globulus* wood meal. Calibrations for NIR prediction of cellulose content indicated that NIR analysis could be used as a reliable predictor (standard errors of predictions (SEP) < 2.5%, with a maximum of 6 principle components (PCs)). The data were converted before analysis to

the second-derivative mode and then analyzed using PLS. This study demonstrated a cost-effective method for screening large numbers of trees for cellulose content based on NIR analysis. (8)

Poke *et al.* (2004) predicted extractives and lignin contents in wood meal of the same species using the same technique. Laboratory measurements were highly correlated with NIR predicted values (coefficients of determination (R^2) of 89% for extractive content, 83% for acid-soluble lignin content, 97% for Klason lignin content (dissolution of carbohydrates in sulfuric acid), and 99% for total lignin content). It was concluded that NIR analysis was a reliable predictor of extractive and lignin content in *E. globulus.* (9)

Baillères *et al.* (2002) employed NIR reflectance analysis as a tool for rapid screening of some major wood characteristics in a *Eucalypt* breeding program. From a narrow genetic base, wood-meal samples (*E. urophylla* x *E. grandis*) with fixed moisture content were analyzed to determine quantitative relations between NIR spectral bands and extractive content and lignin composition. The first key results of this study were that a reproducible spectrum could be obtained and variation in particle size did not have a significant effect on the spectra. The results also revealed that NIRS data that were transformed to second-derivative mode and analyzed by PLSR could be used effectively to predict characteristics linked closely with the chemical composition of wood (SEP < 5%, with a maximum of 10 PCs). The statistical parameters of the calibration equation applied were improved after wood extractives were eliminated from the analysis. The

chemical assignment of the major absorbance bands in the NIR region of the eucalyptus spectrum are shown in **Table 1.1** (10)

Brinkmann *et al.* (2002) compared different methods for lignin determination as a basis for calibration of NIR reflectance spectroscopy. The acid detergent fiber (ADF) and the thioglycolic acid (TGA) lignin data of beech were used to calibrate NIR spectra of dry beech powder for lignin prediction. The calibration equations were calculated using PLSR and pre-processed by applying the second-derivative. Both NIR calibration procedures based on different tissues gave good statistical fits (with correlation coefficients (r) = 99% and SEP < 4%) indicating that TGA and ADF lignin concentrations of beech could be estimated by NIRS with high accuracy. (11)

Terdwongworakul *et al.* (2005) determined the chemical composition of wood meal of *Eucalyptus camaldulensis* by using NIR reflectance spectroscopy. MLR (multiple linear regression) analysis and PLS analysis were introduced to develop the optimum calibration equations for contents of cellulose, pentosans, and lignin in wood with variation of either pre-treatment of NIR spectra or selected wavelengths. In this study, MSC, smoothing, normalization, first- and second-order derivations were employed as the spectrum pre-treatment. In MLR analysis, a reasonably good model was found only for pentosans (r = 90% and SEP = 4%). The PLS analysis improved the accuracy of prediction for every criterion variable, especially for pentosans and lignin. The band at 1676 nm (assigned to the CH stretch in an aromatic skeletal structure) was strongly correlated to lignin. (12)

Table 1.1. Chemical assignment of the major absorbance bands of the eucalyptus NIR spectrum (10)

Wavelength (nm)	Local Mode	Structure
1394	CH combination	CH_2
1520	NH stretch 1st overtone	$CONH_2$
1616	CH stretch 1st overtone	$=CH_2$
1688	CH stretch 1st overtone	Aromatic
1724	CH stretch 1st overtone	CH_2
1740	S-H stretch 1st overtone	-SH
1782	CH stretch 1st overtone	Cellulose
1896	OH stretch, CO stretch	C=O, CO_2H
1910	OH stretch 1st overtone	Aromatic-OH
1992	NH stretch bend combination	Amino acids
2028	C=O stretch 2nd overtone	$CONH_2$
2074	NH_2 deformation 2nd overtone	Amide II
2266	OH CO combination	Cellulose
2280	CH CH_2 deformation combination	CH_3, starch
2296	CH stretch bend 2nd overtone	Protein
2332	CH stretch, CH deformation	Cellulose, starch
2386	CO stretch OH deformation 2nd overtone	Primary alcohols ROH

Poke and Raymond (2006) predicted extractives, lignin, and cellulose contents utilizing NIR reflectance spectroscopy on solid wood in *E. globulus*. Predicting wood traits from NIR spectral data taken from solid wood would greatly increase the speed and cost-effectiveness of this procedure. Existing ground wood calibrations (9) were

evaluated for the prediction of wood chemistry from NIR spectral data taken from solid wood. Extractives, acid-soluble lignin, and Klason lignin contents were poorly predicted. Total lignin and cellulose contents showed moderate relationships between laboratory values and the NIR predicted values. NIR calibrations were further developed specifically for predicting wood chemistry from solid wood. All calibrations had high R^2 values ranging from 72% to 88%, and standard errors of calibrations were less than 1.37%. Calibration validation produced high correlation coefficients between predicted and laboratory values for extractives, Klason lignin, total lignin, and cellulose contents with R^2 values ranging from 67% to 87%. Acid-soluble lignin was poorly predicted. This study showed that NIR analysis on solid wood of *E. globulus* could be reliably used to predict extractives, lignin, and cellulose contents. It also determined that existing ground wood calibrations, although they could give crude estimates of the wood chemistry values, would need to be re-developed for accurate predictions from solid wood. (13)

The ability of reflectance NIRS together with PLSR to predict extractive content of a ground tropical hardwood (*Astronium grveolens*) was investigated by Taylor *et al.* (2008). NIR spectra accurately predicted hot-water-soluble extractive content of this naturally durable wood (root mean square error of prediction (RMSEP) = 19%, with 9 PCs). (14)

Zahri *et al.* (2008) applied reflectance NIRS combined with PLS analysis to quantitatively assess total phenol contents of European oak (*Quercus petraea* and *Q. robur*) on solid heartwood surfaces. The spectra were recorded separately from the

longitudinal-radial and the transverse section surfaces by reflectance. The spectral data were then pre-treated by several pre-processing procedures, such as MSC, first- and second-derivative and standard normal variate (SNV). The tannin contents of sawmill collected from the longitudinal-radial and transverse section surfaces were first determined by quantitative extraction and then, total phenol contents in tannin extracts were measured. The NIR data were correlated against the laboratory results. Calibration models built with PLSR displayed strong correlation and low RMSEP ($R^2 < 82\%$ and RMSEP < 13%, with maximum of 6 PCs). The best calibration was provided with second-derivative spectra (R^2 of 93% for the longitudinal-radial plane and of 91% for the transverse section plane). This study illustrated that the NIRS technique when used in conjunction with multivariate analysis could provide reliable, quick and non-destructive assessment of European oak heartwood extractives. (15)

1.3.1.2. Softwoods

Yeh *et al.* (2004) characterized the lignin content of wood using transmittance NIR analysis. Unlike reflectance, transmittance techniques, which penetrate fully through the sample, are less sensitive to sample preparation and homogeneity and permit the analysis of smaller quantity samples. Using MLR and PLS statistical analysis the lignin contents of dried loblolly pine wood wafers, taken from 12-mm tangential surface of increment cores, and synthetic wood, prepared by blending milled wood lignin and holocellulose, were compared and quantified. The peak at 1672 nm was assigned to lignin and the peaks at 1366, 1432, and 1588 nm to carbohydrates. The band ratios showed a linear

relationship with increasing lignin content. Strong correlations (R^2 close to one) were obtained between the results predicted by NIR and those obtained from traditional chemical methods for both techniques. In addition, NIR results from wood samples with different particle sizes and various lignin contents were discussed. Contrary to previous reports in the literature, they showed that variations in particle size between 30 and 60 mesh did have an effect on the NIR spectra. To eliminate the particle size effect they suggested utilizing the ratio of the components peaks, rather than the individual peak data. (16)

Sykes *et al.* (2005) used the same method and type of samples to develop calibration models for the estimation of cellulose and lignin contents of loblolly pine. The second-derivative NIR spectra (600-1900 nm) showed noise at both ends of the spectra, therefore calibration equations were based on a reduced spectrum (800-1600 nm for cellulose and 800-1750 nm for lignin). NIR calibrations and laboratory measurements based on one site were generally reliable, with R^2 ranging from 55% to 88%. Predicting properties in one annual ring using calibration equations from another ring showed potential for predicting cellulose content with R^2 values of approximately 60%. Predicting the wood properties using the calibration equations from one site to predict another showed moderate success for cellulose content (R^2 = 63%). However, prediction of lignin content was not as reliable, partially because of low variation in lignin content in these wood samples and large errors in measuring lignin content in the laboratory. (17)

Kelley *et al.* (2004) used reflectance NIRS coupled with PLSR to predict the chemical components (lignin, extractives, glucose, xylose, mannose, and galactose) of uniform solid loblolly pine wood, and evaluated the impact of reducing the spectral range from 500-2400 nm to 650-1150 nm. The samples were selected from different radial locations and heights of three loblolly pine trees. Strong correlations between the NIR data and traditional wet chemical techniques were obtained, with r generally above 80%. The correlation coefficients remained high even when the spectral range was reduced. (18)

Yeh *et al.* (2005) used transmittance NIRS to predict the variation in chemical composition of loblolly pine and aspen (hardwood) wood wafers. The effects of sample preparation, sample quantity (single vs. stacked multiple wood wafers), and NIR acquisition time on the quantification of cellulose and lignin content were investigated. Strong correlations were obtained between wet chemistry values and the NIR-predicted values for both species (SEP < 3%, with maximum of 7 PCs). Stronger correlations were obtained for the stacked-wafer model than for the averaged single-wafer model. Moreover, using stacked wafers was not only faster but resulted in a better signal-to-noise ratio, which led to reduction in the calibration error. The calibration error could be reduced by some technical improvements, such as, increasing the signal intensities, increasing the accuracy of the reference method, or broadening the variability of the calibration set. (19)

The same research group (2006) used the same technique and analysis to characterize the chemical components of transgenic products. Two types of sample preparations were

studied and discussed. In the first method, wood meal pellets made from only 75 mg of aspen wood meal, showed very strong correlations between the NIR obtained data and conventional wet chemistry results for the lignin content, syringyl/guaiacyl ratio (S/G), cellulose, and xylose content. The second method, which used three-layer wood wafers, did not produce good correlations. This may be due to the poor intensity of the NIR spectrum as a result of only being able to use three wood wafers per specimen. Thus, transmittance NIRS using wood meal pellets represented a rapid and reliable analytical tool for studying chemical components of wood requiring small sample amounts. (20)

All loblolly pine trees, especially the juvenile portion, contain various amounts of compression wood which is formed when grown under stress. The morphological, chemical, and papermaking properties of compression wood are distinctively different from those of normal juvenile wood and mature wood. Compression wood has higher lignin and galactan, but lower cellulose and mannan content. Chen *et al.* (2007) used NIRS and other micro analytical methods to quantitatively determine the percentage of compression wood in an incremental core of loblolly pine wood meal. Lignin, cellulose, and sugar (glucan, galactan, and mannan) contents were determined using transmittance NIRS. Excellent correlations were obtained between the NIR prediction and laboratory data for all traits ($R^2 > 92\%$). Compression wood content in an increment core could be accessed by plotting galactan, glucan, mannan, or cellulose content against lignin content. These chemical properties could be determined by transmittance NIRS. (21)

Jones *et al.* (2008) compared loblolly pine whole-tree wood property calibrations using reflectance NIR spectra obtained using a variety of sampling options. Calibrations for whole-tree lignin obtained using PLSR on NIR spectra of whole-tree chips (milled or intact) had the strongest statistics; calibrations based on NIR spectra from milled increment cores were similar. Calibrations based on NIR spectra obtained from intact cores provided weaker statistics than those obtained using milled cores. Other options for sampling the tree (drill shavings, collecting spectra from tree in the field, etc.) gave errors that were too large for practical applications. If an increment core is going to be used to estimate whole-tree properties, it was recommended that it be dried and milled prior to analysis. (22)

1.3.2. Physical and mechanical properties of wood

1.3.2.1. Moisture Content

Wood is a typical hygroscopic material with porous structure. It exhibits the characteristic behaviour of swelling accompanied by a remarkable change in mechanical or physical properties when adsorbing or desorbing water. Therefore, the monitoring of MC is very important for wood the industry. Tsuchikawa *et al.* (1996) developed calibration equations for MC with sufficient precision by using NIRS together with MLR analysis. When the wavelengths as the explanatory variables were selected (1272, 1672, and 1960 nm) suitably for the calibration equations, they could predict the MC independently of the ten wood species analyzed (r = 98% and SEP = 1.85%). (23)

Axrup *et al.* (2000) evaluated NIR reflectance measurements with a silicon diode array spectrometer to investigate the determination of water of wood chips. The water content ranged from 19 to 56% and the predictions models with three PCs had good accuracy with RMSEP values less than 3.2%. (7)

Jonsson *et al.* (2004) presented an approach for on-line control and monitoring of pulpwood chip properties based on reflectance NIRS. The pulpwood chips used as raw material in a pulp and paper making process were characterized at- and on-line using NIR spectroscopic measurements. Collected NIR spectra were used in multivariate calibration models for prediction of the MC (35-65%) as well as the between- and within-species variation in the studied raw material. Statistical experimental design was used to form a calibration data set including most of the variation occurring in a 'real' on-line situation. NIR spectra for all designed samples were measured at-line and the estimated calibration models were used for carrying out predictions on-line. Predictions of the MC as well as the percentage contents of pine and sawmill chips in the raw material were carried out using PLSR methodology. To validate the quality of the predictions, wood chips from the studied process were sampled and analysed in the laboratory before being subjected to predictions in the on-line model. Comparison of the filtered on-line predictions with the results obtained from the laboratory measurements indicated that moisture (RMSEP = 3.78%, with 4 PCs) and pine chip contents could be well predicted by the on-line model, while predictions of sawmill chip content showed less promising results. (24)

Reflectance NIRS, coupled with PLSR, was used to predict the green MC and density of solid red oak by Defo *et al.* (2007). NIR spectra were collected from tangential, radial and transverse surfaces of the samples. PLS models were validated using an independent test set. In general, spectra collected from transverse and radial surfaces gave better predictions than the ones collected from tangential surfaces. Good predictions were obtained for spectra collected from transverse or radial surfaces, with RMSEP of less than 3.6% for MC (70-100%) and 19.8 kg/m^3 for basic density when using eight PCs. (25)

Adedipe and Dawson-Andoh (2008) examined the feasibility of using reflectance NIRS combined with multivariate data analysis to predict MC (0.3-80%) of yellow-poplar veneer sheets. All the spectral data were normalized and pre-processed using second-derivative transformation. Both principle component regression (PCR) and PLSR techniques indicated clustering of veneer samples of the same or close MC range with a clear distinction between samples of low and high MC. Predictions had R^2 greater than 94%. The spectral window, 1400-1900 nm, between the two MC peaks (1450 and 1930 nm) gave R^2 of 98.5% and 98.6% for PCR and PLSR, respectively. There was no clear distinction between the PCR and PLS models developed using the NIR spectra region of 1400 to 1940 nm. However, the PLS models with lower RMSEP, SEP and bias were better when compared to the PCR models developed using the whole NIR spectra region. (26)

1.3.2.2. Density and Mechanical Properties

Clearwood is wood that does not contain characteristics such as knots, cross grain, checks, and splits. This wood has anatomical characteristics such as growth rings that occur in consistent patterns. Clearwood specimens are usually considered "homogeneous" in wood mechanics. Wood may be described as an orthotropic material; that is, it has unique and independent mechanical properties in the directions of three mutually perpendicular axes: longitudinal, radial, and tangential. The longitudinal axis is parallel to the fiber (grain); the radial axis is normal to the growth rings (perpendicular to the grain in the radial direction); and the tangential axis is perpendicular to the grain but tangent to the growth rings. Clearwood mechanical properties are a function of density, wood chemistry, and microfibril angle (MFA) and thus, NIRS can be used to measure the stiffness of increment cores with good accuracy. (6)

The optimal use of a given wood raw material in the pulp and paper industry is dictated by the underlying wood physical and mechanical properties. Wood density is considered a key property, affecting, for example, pulp yield per unit of wood volume. Hoffmeyer and Pedersen (1995) investigated wood properties of Norway spruce by reflectance NIRS. The spectral data of cross sections of clearwood were compared to such properties as MC, density, and compression strength using PLSR and PCR. In addition, the spectral data were compared to the bending strength of the structural timber from which it had been cut. The NIR measurements were pre-treated by normalization and derivation. The NIR dependency on surface roughness (circular sawn, band sawn,

and sanded) was investigated and found to be of minor importance. NIR calibrations for MC (MC< 30%), density, and compression strength proved that the NIR technique was an excellent non-destructive method ($R^2 > 90\%$). For the prediction of bending strength of timber NIR was less efficient ($R^2 = 29\%$). However, NIR still contributed to timber strength prediction at the same level as annual ring width, the parameter which is presently visually assessed by timber graders. (27)

Schimleck *et al.* (1999) estimated density of *E. globulus* using reflectance NIRS. The spectral data were converted to the second-derivative mode and analyzed by PLSR. They found that the densities of milled 12-mm increment core samples ranging from 378 to 656 kg/m^3 could be determined with an accuracy of prediction of ca. ±30 kg/m^3. This error compared well with the accuracy of prediction of pilodyn density measurements on similar samples of ca. ±22 kg/m^3. The basic densities of increment cores having relatively low basic densities were consistently overestimated and those having relatively high basic densities were consistently underestimated by the NIRS method. These models were not as successful as those obtained by the previous research. For example, model developed by (27) had accuracy of prediction of 19 kg/m^3. Their model also had a much higher R^2 value compared with the models developed in this study. This might be due to the species analyzed. Norway spruce is softwood and, compared with hardwoods such as *E. globulus*, their anatomy is relatively simple. They concluded that the relatively high error of density estimates indicated that NIRS was unsuitable for the accurate estimation of the density of *E. globulus* increment cores. (28)

Schimleck *et al.* (2001) used reflectance NIRS for the prediction of a range of solid wood properties. A series of *Eucalyptus delegatensis* were characterized in terms of density, wood stiffness (longitudinal modulus of elasticity (E_L) which is defined as the rate of change of strain as a function of stress), MFA (defined as the angle that the helical windings of cellulose chains, within the fiber wall, make with fiber axis), and modulus of rupture (MOR) (defined as the maximum fiber stress). NIR spectra were obtained from the radial-longitudinal face of small strips and used to generate calibrations for the measured mechanical properties. The spectra were converted to the second-derivative mode and analyzed using PLSR. The relationships were good in all cases, with R^2 ranging from 77% for MOR through 90% for E_L to 93% for stick density. It was concluded that appropriately calibrated NIRS could form the basis of a "universal" testing instrument capable of predicting a wide range of product properties from single type of spectrum obtained from the product or from the raw material. (29)

A series of *Pinus radiata* samples were later examined by these scientists (2002) using the methods described in (29). Calibrations developed for measured properties had R^2 ranging from 68% for 100/MFA to 94% for density strip. A good relationship (R^2 = 83%) was also obtained for E_L estimated using data collected by SilviScan-2 (reported as $E_{L(SS)}$). It was also observed that the correlation between E_L and $E_{L(SS)}$ was strong for the *P. radiata* samples (R^2 = 86%). They concluded that a NIR instrument could be calibrated to estimate stiffness of increment cores, based on SilviScan-2 data (SilviScan combines x-ray densitometry, diffractometry and image analysis). They also concluded that calibrated NIRS might provide a rapid, inexpensive method for estimating the wood

properties of large numbers of increment core samples. NIR spectra could be collected from the prepared surface of increment cores and then a calibration applied, such as the $E_{L(SS)}$ calibration developed in this study, to provide pith to bark profiles of a given property or to determine average properties of an increment core. They suggested that the good calibration statistics they obtained for MFA might have been due to the systematic within-tree variation in a range of associated properties, such as cellulose content. (30)

The same research group (2001) described application of the same method and analysis to a large number of mixed species that displayed extremely wide variations in wood chemistry, anatomy, and physical properties. The uniform and solid mixed species samples were characterized in terms of density, wood stiffness, and MFA. The calibrations developed for density and E_L had the highest R^2 and demonstrated that it was possible to develop general calibrations for these important wood properties across a wide range of species. These mixed species calibrations were used to estimate wood properties of two species (*E. delegatensis* and *P. radiata*). The results obtained for density and E_L indicated that mixed species calibrations could be used to rank trees. In practice the extreme variation of samples selected for this study would not be required. It is expected that refinement of calibrations, through sample selection, would provide more accurate prediction of physical properties. (31)

Schimleck et al. (2001) found that some species had very large residuals (NIR-fitted values minus laboratory-determined values) when calibrations for the measured properties were developed (31). Two species that were noted as having large residuals

were known to have high levels of extractives and it was suggested by (31) that the development of calibrations using extracted wood might provide improved results for these species. Therefore, Schimleck *et al.* (2003) applied reflectance NIRS and PLSR to extracted wood species representing a diverse array of taxonomic affiliations, wood chemistry and physical properties. The extracted samples were characterized in terms of density, MFA and E_L. NIR spectra were obtained from the radial longitudinal face of each sample and used to generate calibrations for the measured physical properties. Extraction was found to improve the calibration statistics for all properties. (32)

Schimleck et al. (2002) utilized reflectance NIRS for predicting the radial variation of the E_L of increment cores. Sets of dried clearwood samples from *E .delegatensis* and *P .radiata* were characterized in terms of $E_{L(SS)}$. NIR spectra, obtained from the radial–longitudinal face of each sample, were used to develop $E_{L(SS)}$ calibrations for the *E. delegatensis* and *P. radiata* sample sets and the two sets combined. The relationships between laboratory-determined $E_{L(SS)}$ and NIR-fitted $E_{L(SS)}$ were good in all cases. $E_{L(SS)}$ was estimated in separate test sets and found to correlate well with measured E_L. NIR spectra were obtained in 15-mm sections from the radial–longitudinal face of two intact *P. radiata* increment cores. $E_{L(SS)}$ of each section was estimated using the *P. radiata* and the combined *P. radiata* and *E. delegatensis* calibrations. NIR estimates of $E_{L(SS)}$ were in good agreement with SilviScan-2 determined stiffness indicating that NIRS could be successfully used to estimate radial variation in wood stiffness of increment cores (R^2 ranging from 89% to 93%). (33)

Schimleck and Evans (2002) suggested that estimates of the $E_{L(SS)}$ of increment core samples could be improved by using calibration samples that are more closely related to the samples whose properties would be predicted. Moreover, the resolution could be improved by using a smaller window when the NIR spectra were collected. Thus, they selected *P. radiata* increment core samples for the development of a $E_{L(SS)}$ (34) and MFA (35) calibration based on NIR spectra obtained from the radial–longitudinal face of each sample in 10-mm increments. The primary aim of their work in (34) was to investigate if a $E_{L(SS)}$ calibration developed using a subsample of cores representative of a larger set provided better predictions of the $E_{L(SS)}$ than those reported in (33). The $E_{L(SS)}$ calibration was developed using eight factors giving an excellent relationship between SilviScan-2 determined $E_{L(SS)}$ and NIR fitted the $E_{L(SS)}$ (R^2 = 97%) and a low standard error of calibration (SEC). To test $E_{L(SS)}$ calibration, NIR spectra were obtained in 10-mm sections from the radial–longitudinal face of two intact *P. radiata* increment cores and the $E_{L(SS)}$ of each section predicted. NIR estimates of $E_{L(SS)}$ were in excellent agreement with $E_{L(SS)}$ determined using SilviScan-2 data, with R^2 of 99% (core A) and 98% (core B). The predictions were superior to those reported in (33). (34) The MFA calibration was developed using seven factors giving an excellent relationship between SilviScan-2 determined MFA and NIR fitted MFA (R^2 = 95%). NIR predicted MFA was found to be in excellent agreement with MFA determined by SilviScan-2, with R^2 of 98% (core A) and 96% (core B). This data set had a much wider range than the limited ranges in (29, 30). They also investigated the possibility that systematic within-tree variation in a range of associated properties, such as cellulose content was the reason for good calibrations (30). The correlation of MFA with the scores of PLS factors used to develop

the current calibration was examined. The corresponding loadings plot had large loadings in regions of the NIR spectrum assigned to cellulose suggesting that variation in cellulose content is an important factor in the success of the MFA calibration. It was also found that several wavelengths assigned to cellulose had strong correlations with MFA providing further evidence that MFA variation is linked to variation in cellulose content. (35) Both predictions closely followed the patterns of MFA and $E_{L(SS)}$ radial variation determined by SilviScan-2. They concluded that NIRS provided a rapid method for determining MFA and $E_{L(SS)}$ variation in increment cores and was suitable for the routine analysis of large numbers of samples. (34, 35)

Potentially, NIRS could be used to test trees in the field, but it was not known what impact the variable and high MC of standing trees would have on the calibrations. Therefore, Schimleck's laboratory (2003) applied reflectance NIRS and modified PLSR to green loblolly pine radial samples to develop calibrations for the prediction of physical and mechanical wood properties (air-dry density, MFA, and stiffness). NIR spectra were obtained in 10-mm steps from the radial longitudinal and transverse faces of each sample and used to develop calibrations for each property. NIR spectra were collected when the wood was green (MC ranged from 100 to 154%) and dried to 7% MC. Relationship between measured and NIR estimates for green wood were good; R^2 ranged from 0.79 (MFA) to 0.85 (density). Differences between calibrations developed using the radial longitudinal and transverse faces were small. Calibrations were tested on an independent set. Predictive errors were relatively large for some green samples and relationships were moderate; R^2 ranged from 0.67 (MFA) to 0.81 (stiffness). Dry wood calibrations

demonstrated strong predictive relationships with R^2 ranging from 0.87 (density) to 0.95 (stiffness). (36)

Kelley *et al.* (2004) used reflectance NIRS to predict the mechanical properties (MFA, MOR, and modulus of elasticity (MOE)) of oven-dried solid loblolly pine wood. The samples were selected from different radial locations and heights of three loblolly pine trees. These mechanical properties were correlated with the NIR spectra using PLS models. The correlations were very strong, with r generally above 0.80. The mechanical properties could also be predicted using a reduced spectral range (650 nm–1150 nm) that should allow for field measurements of these properties using handheld NIR spectrometers. (18)

The same research group (2004) used the same technique for predicting the mechanical properties (MOE and MOR) of a number of different softwood species. The Vis and NIR (500–2400 nm) spectra and mechanical properties of almost 1000 small clearwood samples with fixed MC from six softwood species: loblolly pine, longleaf pine, slash pine, shortleaf pine, ponderosa pine, and Douglas fir were measured. PLS modeling showed that the NIR spectra of these softwoods could be used to predict the mechanical properties of the clearwood samples. The correlation coefficients for most of these models were greater than 0.80. All six softwood species were combined into one data set and a PLS model was constructed that effectively predicted the strength properties of any of the individual softwoods. Reducing the spectral range to between 650 and 1050 nm only caused a slight decrease in the quality of the models. (37)

When a radial strip is cut from an increment core it can potentially be cut with two orientations: parallel to longitudinal tracheids, i.e., a radial-longitudinal (RL) strip (the orientation of SilviScan samples), or at an orientation of 90° to the tracheids, i.e., a radial-transverse (RT) strip. Both strips could be used for NIR analysis, but it is unknown how calibrations based on RT-face NIR spectra compare with those based on RL-face NIR spectra. Therefore, Schimleck et al. (2005) characterized *P. taeda* RL strips in terms of air-dry density, MFA, stiffness and several tracheid morphological characteristics. Reflectance NIR spectra were collected in 10-mm increments from the RL and RT faces of matching strips with fixed MC and used to develop calibrations for each property. In general, RL-face NIR spectra gave calibrations that provided stronger relationships. Differences between the two sets of calibrations were small, indicating that either face could be used for NIR analysis. (38)

Preliminary studies based on small sample sets showed that NIRS has the potential for rapidly estimating many important wood properties. However, if NIR is to be used operationally, then calibrations using several hundred samples from a wide variety of growing conditions need to be developed and their performance to be tested on samples from new populations. In the study of Jones *et al.* (2005), *P. taeda* radial strips (cut from increment cores) representing 15 different sites from three physiographic regions in Georgia (USA) were characterized in terms of air-dry density, MFA, and stiffness. Reflectance NIR spectra were collected in 10-mm increments from the radial longitudinal surface of each strip and split into calibration and prediction sets. The spectra were treated with various transformations. PLS Calibrations were developed using untreated

and mathematically treated (first- and second-derivative and MSC) spectra. Strong correlations were obtained for all properties, the strongest R^2 values being 0.83 (density), 0.90 (MFA), and 0.93 (stiffness). When applied to the test set, good relationships were obtained (R^2 ranged from 0.80 to 0.90), but the accuracy of predictions varied depending on mathematical treatment. The addition of a small number of cores from the prediction set (one core per new site) to the calibration set improved the accuracy of predictions and importantly minimized the differences obtained with the various mathematical treatments. These results suggested that density, MFA, and stiffness could be estimated by NIR with sufficient accuracy to be used in operational settings. (39)

The estimation of density, MOE, and MOR of loblolly pine clearwood samples from a diverse range of sites across the southern USA was investigated by Schimleck *et al.* (2005) using reflectance NIRS. NIR spectra were obtained from the radial and cross sectional (original, rough, and sanded) surfaces of blocks cut from the ends of juvenile and mature wood samples. In this study, the tangential surface was not examined because it does not represent all of the wood property variation present in a short clearwood sample. Calibrations based only on juvenile or mature wood samples had weak calibration statistics and failed to perform well when applied to a separate test set possibly due to small variability in wood properties. Calibrations developed using both juvenile and mature wood NIR spectra provided good relationship for all properties with R^2 ranging from 0.82 (MOE, radial face) to 0.90 (density, radial face) demonstrating that it was possible to obtain multi-site calibrations for these properties. Prediction R^2 ranged from 0.77 (MOE, radial face and density, original cross section) to 0.86 (MOR, sanded

cross section). Though differences between surfaces were small, on average the sanded cross-sectional surface provided the best calibration and prediction statistics. (40)

Until now, NIRS has been used to predict mechanical properties of solid wood from spectra collected from either solid or grounded wood. In the study of André *et al.* (2006), three yellow poplar wood beam samples were tested in bending (the applied load was increased) while collecting NIR reflectance spectra on their compression and tension faces. The data were then converted to absorbance, normalized, and submitted to MSC to remove scatter due to surface roughness. It was demonstrated that NIRS coupled to PLS analyses could be used to predict tension or compression loads within these wood samples. Strong correlations ($r > 0.96$) between the measured load and the predicted load were obtained using spectra taken from both the tension and compression surfaces of the small wood beams. These strong correlations can provide experts with a new tool to monitor and diagnose timber structures. The regression coefficients obtained from the PLS model for the tension face (all samples together) were different than the ones obtained from the PLS model for the compression face. A parallel was drawn for the compression face between the main compounds identified from the regression coefficients and those involved in the reinforced concrete theory. Yet, it was not possible to draw the same parallel for the tension face due to the presence of chemical groups belonging to both cellulose and lignin. (41)

Thirteen wood parameters were predicted using reflectance NIRS combined with PLSR by Lestander *et al.* (2008). This analysis was done on clearwood samples of Scots

pine from trees at two sites. NIR measurements were conducted in the radial direction on the plane perpendicular to the longitudinal direction. Calibrations based on the measured parameters at seven growth rings (cambial age ranging between 6 and 42 years) could be divided into three groups: (i) the best accuracy was found for E_L ($r > 0.9$) followed by bending, compression, and cell length ($0.8 < r < 0.9$); (ii) MFA, longitudinal hardness, proportion of latewood, and creep with correlations in the range of 0.7-0.8; and (iii) tangential hardness, cell diameter, and cell wall thickness with $0.4 < r < 0.7$. It was also shown that juvenile (cambial age $\leq$ 20 years) and mature wood could be classified using NIR techniques. (42)

Estimation of the density along with the tensile strength of uniform and solid wood within both the elastic and plastic deformation ranges, represented as MOE and ultimate tensile stress (UTS), respectively, were performed using reflectance NIRS by Tsuchikawa *et al.* (2005). Various pre-treatments were performed on the NIR data. A PLS analysis was applied to the measurements of density, MOE, and UTS, and resulted in a high accuracy of prediction, independent of wood species. The correlation coefficient between the criterion variables and explanatory variables indicated that in the case of Japanese red pine, the increase of OH absorption bands for the amorphous regions of cellulose was related to an increase in density. Conversely, the increase of OH absorption bands for the crystalline regions and the decrease of OH absorption bands for the amorphous regions of cellulose were related to an increase of the MOE or UTS. In the case of Japanese zelkova (hardwood), the increase of the crystalline regions in cellulose and the decrease of the amorphous regions in cellulose, hemicellulose, and lignin were related to an increase of

both density and the MOE or UTS. Variation of the regression vector indicated that in the case of density for Japanese red pine, there was a slightly significant feature at the absorption band due to the amorphous regions of cellulose, whereas in the case of Japanese zelkova, there was no dominant absorption range that contributed to the PLS analysis. There were some characteristic peaks in the regression vector for the MOE of the Japanese red pine, which were due to OH absorption bands of cellulose. Japanese zelkova presented different aspects, such as the contribution of the CH absorption band of the semi-crystalline or crystalline regions of cellulose, and the CH absorption band of hemicellulose and lignin, to the validation of the MOE. Finally, they proposed that the key spectroscopic factors for the validation of physical and mechanical properties of softwood, from a perspective of anatomical observation, were absorption bands due to intramolecular hydrogen-bonded OH groups in the crystalline regions of cellulose, which are oriented preferentially in a direction parallel to the cellulose chain, and might strongly affect the tensile strength because of an approximate coincidence between the direction of cellulose chains in the S_2 layer (i.e., the middle layer of the secondary wall) within the cell wall and the tensile direction. Since hardwoods have much more complex structures than softwoods, the most important key factor governing the tensile strength may not be the nature of the cellulose chains, but the interaction between the three principal constituents of wood. They also assigned absorption bands in the second-derivative NIR spectra (**Table 1.2**). (43)

Table 1.2. Chemical assignment of the major absorbance bands of the Japanese red pine and Japanese zelkova NIR spectrum (43)

Wavelength (nm)	Local Mode	Structure
1428	OH stretch 1st overtone	Amorphous regions in cellulose
1476	OH stretch 1st overtone	Semi-crystalline regions in cellulose
1548, 1592	OH stretch 1st overtone	Crystalline regions in cellulose
1672	CH stretch 1st overtone	Aromatic groups in lignin
1724	CH stretch 1st overtone	Furanose/pyranose due to hemicellulose
1790	CH stretch 1st overtone	Semi-crystalline or crystalline regions in cellulose
1830	OH stretch + 2xCO stretch	Semi-crystalline or crystalline regions in cellulose
1916, 1980	OH stretch + OH deformation	Water
2080	OH stretch + CH deformation	Semi-crystalline or crystalline regions in cellulose
2140	CH stretch + CC deformation	Amorphous regions in cellulose
2200	CH stretch + CO stretch	CHO
2272	CH_2 stretch + CH_2 deformation	?
2336	CH stretch + CH deformation	Semi-crystalline or crystalline regions in cellulose
2380	CH deformation 2nd overtone	Holocellulose

It has been shown that NIRS offered a rapid method for the estimation of MFA and $E_{L(SS)}$. Schimleck et al. (2005) suggested that the success of these NIR calibrations might be related to air-dry density, because density varies in wood simultaneously with MFA and stiffness. The importance of density variation was investigated by developing calibrations for MFA and $E_{L(SS)}$ using *P. radiata* and *P. taeda* sample sets where the density range was small and the relationships between density and MFA and density and $E_{L(SS)}$ were poor. Excellent PLS calibrations for MFA and $E_{L(SS)}$ were obtained,

demonstrating that reflectance NIRS could provide strong relationships for MFA and stiffness even when density variation was limited. They suggested that the inferior calibration statistics observed for *P. taeda* sets having densities less than 500 kg/m^3 were probably a consequence of high average MFA for these sets and limited MFA range. As MFA increases, the signal-to-noise ratio in the diffraction patterns decreases, reducing the precision with which MFA can be determined. This reduction in precision contributes to the weakness in the calibrations at high MFA. Based on these observations it is probable that calibrations developed using juvenile wood, which inherently has large MFA would fail to give strong statistics. Examination of loading plots from the MFA and $E_{L(SS)}$ calibrations indicated that variation in wood components such as cellulose, lignin and possibly hemicellulose was important. (44)

Fujimoto *et al.* (2007) investigated the feasibility of reflectance NIRS for estimating MOE and MOR in bending tests. NIR measurements were performed on radial-longitudinal surfaces, where the spectral data were obtained from both the tension and compression surfaces of the bending samples. Two uniform and solid hybrid larch sample sets having large and limited density variation were prepared to examine the effects of wood density on estimation of MOE and MOR by the NIR technique. PLS analysis was employed and it was found that the relationship between laboratory-measured and NIR-predicted values was good in the case of sample sets having large density variation. MOE could be estimated even when density variation was limited. The difficulty of estimating MOR might arise from the fact that the density of the wood sample strongly affects MOR rather than MOE. It was concluded that absorption bands due to the OH group in the

semi-crystalline or crystalline regions of cellulose strongly influenced the calibration for bending stiffness and strength of hybrid larch. This was also suggested from the result that both cellulose content and cellulose crystallinity showed moderate positive correlation to wood stiffness. (45)

The composite structure of the S_2 layer in the wood cell wall is defined by the angle of the cellulose microfibrils and concentration of polymers and this structure impacts strength and stiffness. Via *et al.* (2009) used absorbance NIRS and X-ray diffraction to determine the effect of lignin and cellulose associated wavelengths, MFA, density, and radial position within the tree on strength and stiffness. The aromatic portion of lignin provided a good predictive role on strength and stiffness at high MFAs. However, in mature wood where MFA and lignin content was low, cellulose associated wavelengths became increasingly important. The increased importance of the aromatic portion of lignin (1665 nm) on the strength as MFA increased was attributable to the plastic deformation of lignin that occurred beyond the yield point. Finally, a fourfold increase in stiffness was observed when the microfibril angle dropped from 40 to 5°. (46)

1.3.2.3. Anatomical Parameters

Hauksson *et al.* (2001) investigated the feasibility of using reflectance NIRS and PLSR to characterize the basic wood properties of Norway spruce. The discs cut from 40% tree height were used (i.e., where the largest variations in annual ring widths and wood density were found). These discs were used for measuring annual ring widths,

wood density, average fiber length, and the fiber length distributions. The wood meal prepared from these discs, which was dried to ensure uniform humidity, was scanned by NIR and analyzed by PLSR. The results presented in this study demonstrated good to excellent correlation between NIR spectra of wood meal and a number of wood properties. It was even possible to correlate fiber length distributions with these wood properties. The use of a pre-processing method called orthogonal signal correction (OSC) greatly improved the modeling power of the PLS models based on NIR spectra and even fiber length distributions. OSC algorithm removes structure from the NIR spectra which is unrelated to the y variables being modeled. (47)

Estimation of tracheid length by (47) had promising results but they used milled wood samples. Schimleck et al. (2004) employed reflectance NIRS for predicting tracheid (fiber) length of solid loblolly pine wood samples. The 10-mm sections of radial strips were selected and NIR spectra were obtained from the radial longitudinal face of each section. The fibers in these sections were characterized in terms of arithmetic and length-weighted mean tracheid length using a fiber quality analyzer, and calibrations with NIR spectra were developed for both measures of tracheid length by PLSR. Relationships were good, with R^2 of 0.88 for arithmetic tracheid length and 0.96 for length-weighted tracheid length. The accuracy of NIR predicted length-weighted tracheid length was sufficient for ranking purposes. (48)

Schimleck and Evans (2004) selected *P. radiata* increment core samples for the development of calibrations for several tracheid morphological characteristics:

coarseness, perimeter, radial and tangential diameter and wall thickness. NIR spectra, obtained from the radial-longitudinal face of each core in 10-mm sections from pith to bark, were used to develop the calibrations by PLSR. Calibrations for coarseness and wall thickness were excellent, R^2 of 0.91 and 0.89, respectively. Calibrations for the remaining characteristics were weaker (R^2 ranged from 0.65 to 0.69). To test the predictive ability of the calibrations, two intact *P. radiata* increment cores (core A and B) were selected from the same set as the calibration samples. NIR-predicted tracheid coarseness and wall thickness were in strong agreement with measured (SilviScan-determined) values. Radial patterns of variation (NIR-predicted, measured) closely followed each other for both cores, but coarseness and wall thickness were underestimated for core B. Tracheid tangential perimeter was well predicted with R^2 of 0.69 (core A) and 0.79 (core B). Relationships for the remaining characteristics were weak. They suggested that collection of NIR spectra in smaller increments, to capture more of the variation, could improve calibration. (49)

Sykes *et al.* (2005) used transmittance NIRS and PLSR to characterize physical properties of thin wood wafers cut from 12-mm increment cores to develop calibration models for the estimation of average fiber length and fiber coarseness. NIR calibrations and laboratory measurements based on one site were generally reliable. Predicting one ring properties using other ring calibration equations showed potential for predicting fiber coarseness, with R^2 values of ca. 0.60, indicating the potential for early selection. Predicting the wood properties using the calibration equations from one site to predict

another showed moderate success for fiber coarseness ($R^2 = 0.63$), but predictions for fiber length were relatively poor ($R^2 = 0.43$). (17)

NIR wavelengths can provide either useful or misleading calibrations depending on the context. This can happen since a property such as tracheid length is not directly related to the absorbance at any wavelength but is instead the result of a secondary correlation with some unknown chemical constituent. Via *et al.* (2005) investigated the effect of tree age and height on reflectance NIR predictability since tracheid length and chemistry might vary as a function of location within the tree. The NIR data were transformed to first-derivative mode and analyzed by PLSR. It was found that tracheid length predictability of solid longleaf pine did not change with height but decreased with age. As a result, predicting tracheid length regardless of age was good ($R^2 = 0.72$) while predictability with age and height held constant was mostly low to moderate with the exception of rings one and four where it was reasonably good. The modeling of tracheid length, for tree improvement applications needs more justification, particularly for mature wood zones. Without the variation in xylem age, NIR does not generally have an acceptable precision to estimate tracheid length. The exception to this was the rings adjacent to the pith. As a result, NIRS might be suitable for tree improvement in this juvenile zone. (50)

1.3.3. Wood Classification

Knowing the wood species of a board is generally very useful. Lumber can be sorted by species or species group for different end uses and then sold for different prices. For example, higher grades of Douglas-fir typically sell for more than the same grade of SPF lumber. Species differences are also important in lumber-manufacturing processes. For example, dry-kiln efficiency can be improved by putting species with similar drying rates in the same kiln charge. Sorting by species is currently performed by eye using appearance criteria and is quite a difficult task with some species. An automated method for quickly and accurately identifying wood species would relieve graders from a demanding and often difficult task, while increasing product quality and decreasing production costs. (51)

An experiment in identifying the solid wood species of Douglas-fir, lodgepole pine, and Engelmann spruce wood samples was described by Brunner *et al.* (1996), based on their spectral-reflectance curves in the Vis and NIR regions and quadratic discriminant analysis. The direction of illumination was parallel to wood fiber direction to minimize specular reflection. The measurements were made on 25 samples with the same thickness for each species at 5-10 random locations in the sample and taken every 5 nm from 400 to 1100 nm, but only every tenth reading was used in the discriminant analysis. Species predictions based on this method sometimes confused spruce with pine, but achieved an overall accuracy of 98.7% in separating Douglas-fir from spruce/pine. They concluded that further experiments with considerably larger sample sizes would be needed to

confirm this level of accuracy. Improvements in the identification model are also possible. Determining if heartwood or sapwood spectra are better predictors and a further partition of reflectance spectra into earlywood and latewood classes might improve accuracy. (51)

The possibility to make species predictions of mixtures of wood chips from three different wood species (Swedish pine, Swedish spruce and Polish pine) was investigated by Antti *et al.* (1996) based on NIRS. Mixture design and PLS were used for the multivariate calibration modeling. The calculated model was validated both internally and with an external test set. They obtained a PLS model with the overall predictive ability, $R^2 = 0.91$ according to cross-validation and good prediction of the test set objects, $R^2 = 0.78$. The fact that the model gave such good predictions of the test samples was confirmation that this methodology could be used for predicting mixtures of different wood species as well as within species variations. However, samples were finely ground for mixing prior to analysis which is not practical for the applications investigated in this study. (52)

Lebow *et al.* (1996) used spectral reflectance to classify wood-surface features on 40 Douglas-fir veneer samples. The feature types they considered were sapwood-earlywood, sapwood-latewood, loose and tight knots, pitch pockets and streaks, and wane (cambial interface). They measured the spectral reflectance of each feature in 10 nm increments from 400 to 1100 nm. They used both linear and quadratic discriminant analysis to

classify the features and found that quadratic discriminat analysis proved significantly superior, achieving an overall misclassification rate of about 1-2%. (53)

Previous research showed that a variety of surface features on Douglas-fir veneer could be accurately classified via Vis-NIR region reflectance spectra and quadratic discriminant analysis. A paper from the same laboratory (2001) extended those results to an expanded set of features, a broader spectrum and a larger set of physical samples. The wood analyzed came from nine types of surface features (heartwood, sapwood, heart/sap mix, loose and tight knots, pitch pockets and streaks, white speck, and nonfungal stain). The data were collected in the same way as the previous research. They also tested two methods for eliminating the classification procedure's reliance on raw spectral-reflectance curves. Instead of working with the raw curves, which were difficult to obtain by traditional means at sufficient speeds in a production environment, the data required by these two methods (one derived from PCA (principle component analysis) and the other from a simulated extended-color system that used bandpass filters) was much reduced and could potentially be obtained from a video camera equipped with either custom or commercially available bandpass filters. The paper showed that classification accuracies achieved with either of the two reduced-data methods were comparable to the accuracies achieved when using raw spectral data. (54)

Tsuchikawa *et al.* (2003) dealt with a new non-destructive discriminant analysis by which solid wood could be classified on the basis of a combination of absorbance NIRS and Mahalanobis' generalized distance (a measure of generalized distance between

samples based on the means, variances, and covariances of various properties of replicate samples in multivariate analysis). Its accuracy and reasonability were examined for nine wood species with various MCs ranging from oven-dried to a fully saturated free water state. In a discriminant analysis employing second-derivative spectra, each wood group was well distinguished where the selected wavelengths were often derived from the absorption of wood components. When the wavelengths for analysis were limited to the short range of 800–1100 nm, the correct results diminished. In this case, the wavelength derived from the absorption of water (around 980 nm) was often selected. Hence, there may be two bases for selecting the wavelength: (1) when the chemical component of wood substance relates to the discriminant analysis; and (2) when the difference in MC with wood species relates to them. Mahalanobis' generalized distances between softwoods were relatively independent of analytical pattern so it was difficult to explain the difference in their spectroscopic characteristics, whereas the distances between hardwoods were large for easy classification. (55)

The same research group (2003) used the same technique and several chemometric analyses to distinguish various wood-based materials (solid wood, laminated wood, particle- or fiberboard, impregnated wood, and overlaid wood). The concept of Mahalanobis' generalized distance, K nearest neighbours (KNN), and soft independent modeling of class analogy (SIMCA) were evaluated to determine the best analytical procedure. These techniques are used for classifying objects based on distances. The differences in the accuracy of classification with the spectrophotometer, the wavelength range as the explanatory variables, and the light-exposure condition of the sample were

examined in detail. It was difficult to apply Mahalanobis' generalized distances to the classification of wood-based materials where NIR spectra varied widely within the sample category. The performance of KNN in the NIR region (800–2500 nm), for which the device used in the laboratory was employed, exhibited a high rate of correct answers of validation (> 98%) independent of the light-exposure conditions of the sample. When employing the device used in the field, both KNN and SIMCA revealed correct answers of validation (> 88%) at wavelengths of 550–1010 nm. These results suggested the applicability of NIRS to a reasonable classification of used wood at the factory and at job sites. (56)

Jonsson et al. (2004) presented an approach for on-line control and monitoring of pulpwood chip properties based on NIRS and multivariate data analysis. NIR spectra for all designed samples were measured at-line and the estimated calibration models were used for carrying out predictions on-line. Predictions of the percentage contents of pine and sawmill chips in the raw material were carried out using PLS methodology. Comparison of the filtered on-line predictions with the results obtained from the laboratory measurements indicated that pine chip contents could be well predicted by the on-line model (RMSEP= 22%), while predictions of sawmill chip content showed less promising results. (24)

Considerable areas in the northern parts of Norway are afforested with Norway spruce, Lutz spruce, and Sitka spruce. The species have different machining and wood properties but are similar in visual appearance. Flaete *et al.* (2006) evaluated whether

reflectance NIRS combined with multivariate statistical modelling could be used to identify wood from these three species. In all, 83 wood specimens were available for analyses, 36 of which were used as a test set for model validation. NIR spectra were obtained on the air-dry cross-sectional surfaces. The faces were prepared by circular sawing. The reflectance data expressed as apparent absorbance were centered and transformed to first-derivative. An initial PCA indicated that little information from the first and second components could be used for discrimination, but in score-plots of the third and fourth components the samples from the tree species formed clusters. This showed that the NIR spectra did contain information relevant for tree species identification, and that only a small fraction of the total variance could be used for that purpose. For classification of the wood specimens, PLS discriminant analyses were applied. All 47 specimens in the training set were fitted into the correct group. The test set validated results showed that except for two wood specimens, all specimens were correctly classified. The two misclassified samples were Sitka spruce. This is not surprising as the regression coefficients for Norway spruce and Lutz spruce showed a strong inverse relationship along the whole spectral range. Conversely, there were only narrow spectral areas where the PLSR coefficients for Sitka spruce were high and the corresponding regression coefficients for Norway spruce and Lutz spruce were close to zero. Despite the small calibration set and the structure of regression coefficients, the PLS discriminant analysis model performed well also for Sitka spruce. This underpins the potential for calibrating PLS models with higher classification performance if a large number of calibration samples are used. (57)

1.3.4. Treated Wood

1.3.4.1. Chemically and Preservative Treated Wood

There is a growing need to find a rapid, inexpensive, and reliable method to distinguish between treated and untreated waste wood. So *et al.* (2003/4) used NIRS together with multivariate analysis to identify water and oil-borne preservatives. Three types of water-borne preservatives were chosen: CCA (chromated copper arsenate), ACZA (ammoniacal copper zinc arsenate), and ACQ (alkaline copper quat). Several commercially treated deckboards were obtained for this study: CCA/Hemlock-Fir, CCA/Eastern Hemlock, ACZA/Douglas Fir, and ACQ-Hemlock-Fir. They were milled and scanned with a NIR spectrometer. MSC was applied to the spectra prior to multivariate analysis. The spectral data were then analyzed by PCA. The samples separated out according to treatment and were easily distinguishable. Furthermore, separation according to wood species and assay zone was also observed. Within the range of preservative concentrations available, PLSR was also performed on the NIR data, from which retention levels were well predicted. Two oil-based preservatives were also studied: copper napthenate and oxine copper. Again these samples separated and clustered according to their treatments. Furthermore, with the copper napthenate samples, there was a clear graduation in solution strength. The results clearly demonstrated that this technique had potential for use in a variety of recycling and sorting applications. (58, 59)

As with all preservative systems, efficacy is dependent upon reaching a preservative concentration in the treated wood that is above the toxic threshold of the target pest(s), and thus it is important to be able to accurately determine preservative concentration. Reflectance NIRS coupled with PLSR was assessed by Taylor and Lloyd (2007) as a possible method for determining quantitatively boron-based preservative concentration in treated wood. It was found that NIR-based model could successfully predict the borate concentration of treated southern pine sapwood cubes ($R^2 = 0.86$), suggesting that it might provide the basis for rapid, easy-to-operate and portable analytical tool. (60)

Reflectance NIR spectral analysis was used by Peydecastaing *et al.* (2006) to determine the degree of substitution (DS) of chemically modified cellulose and lignocellulosics. Two kinds of samples were studied, long aliphatic-chain cellulose esters and Scots pine wood sawdust chemically-modified either by anhydrides or by ethylene carbonate. It was possible to determine the DS of such samples through a correlation by PLS of the second-order derivatives of NIR spectra. This technique was efficient even when DS values were low, which is difficult to achieve by using FT-IR. It was also possible to distinguish reagent molecules that were attached to the cellulosic substrate by hydrogen bonding from those linked by covalent bonding. (61)

Impregnation of solid wood with a furfuryl alcohol monomer that is subsequently cured using heat and a catalyst to yield an inert polymer inside the cell wall structure provides positive changes in wood properties. In general, these properties are positively correlated with increasing polymer mass, as evaluated by the weight percent gain. Venås

and Riińnan (2008) used NIRS to determine the degree of treatment in Scots pine solid wood modified with *in situ* cured furfuryl alcohol. The treatment range was approximately 0 to 30% weight gain relative to wood dry mass. Reflectance NIR spectra were recorded directly on wood surfaces without sample preparation. PLSRs were used to construct a model for weight percent gain; initially for one data set but later extended to an external test set not treated fully in the same way as the original set. This gave rise to an increase in the prediction error which was neutralised by addition of a small fraction of the new samples in the original calibration set. Four spectral ranges and 13 pre-processing methods were tested and the final model with addition of 30% of the new samples in the original calibration set resulted in a model with a RMSEP of 1.7±0.1% weight gain and a R^2 of 0.97±0.01. The range 1100–2498 nm and by using SNV with second-derivative gave the lowest prediction errors, and this was consequently chosen as the optimal wavelength range and pre-processing method for the prediction of weight percent gain in Scots pine solid wood by NIR. (62)

A new optical system was developed and applied to automated separation of wood wastes, using a combined technique of Vis–NIR imaging analysis and chemometrics by Kobori *et al.* (2008). Three kinds of typical wood wastes were used, i.e., non-treated, impregnated, and plastic-film overlaid wood. The classification model based on SIMCA was examined using the difference in luminance brightness of a sample. Their newly developed system showed good/promising performance in separation of wood wastes, with an average rate of correct separation of 89%. Hence, it was concluded that the

system was efficiently feasible for online monitoring and separation of wood wastes in recycling mills. (63)

1.3.4.2. Thermally Treated Wood

Schwanninger *et al.* (2004) applied reflectance FT-NIRS to thermally modified wood. The NIR measurements were made on a cross-sectional surface of each solid wood sample. The spectral data were then transformed to second-derivatives and analyzed by PCA. A comparison of NIR spectra of untreated beech with wood treated at 220°C showed changes for OH, aromatic CH, CH_3, CH_2, and CH overtone vibrations due to e.g. lignin and hemicelluloses degradation. A close relationship between the chemical changes and the NIR spectra of thermally modified beech wood made clustering easy and provided a possibility for quality control. (64)

The paper by Mitsui *et al.* (2008) dealt with the evaluation of thermally treated Sitka spruce wood by NIRS. The wood samples were treated by steaming at 140°C for 5, 10, 20, 50, and 100 hours. In the NIR second-derivative spectrum, the absorption band at 1447 nm appeared with the procession of heat treatment, which conclusively assigned to the phenolic hydroxyl groups due to the lignin in comparison with the spectrum of acetylated wood. As a result of the changes in the ratio of the areal integral calculated from spectral separation in the region of hydroxyl groups (1389-1639 nm) by the Gauss-Newton method, it was clear that the degradation of hydroxyl group in the cellulose started predominantly from the amorphous region and followed to semicrystalline and

crystalline region. There was an obvious correlation between the weight decrement of wood and the decrement of hydroxyl groups in the cellulose by heat treatment. (65)

NIRS was tested by Esteves and Pereira (2008) for predicting the properties of heat treated wood using pine (*Pinus pinaster*) and eucalypt (*E. globulus)* woods with two types of treatment: in oven and in a steam autoclave. Mass loss, equilibrium MC, dimensional stability, MOE, bending strength, colour CIELAB parameters and extractives content were determined. NIR spectra were obtained using a fiber probe on the radial surface of the samples. PLS models for mass loss showed very high coefficients of determination ranging from 96–98%. The models obtained for wood properties were in general good with coefficients of determination ranging from 78–95% for equilibrium MC, 53–78% for dimensional stability, 47–89% for MOE, 75–77% for bending strength and 84–99%, 52–96% and 66–98% for colour parameters *L*, *a** and *b**, respectively. Coefficients of determination of the models for extractive content varied between 41.9–79.8% for pine and between 35.3–82.2% for eucalypt wood. NIRS showed a good potential for quality control and characterization of heat treated woods. (66)

1.3.5. Weathering, Aging, and Biodegradation of Wood

1.3.5.1 Biodegradation of Wood

Decay and discoloration caused by fungi are major sources of loss in value for both timber production and wood in service. A better understanding of the chemical parameters involved in brown-rot decay is a necessary prerequisite for developing

methods of protecting wood from brown-rot degradation in a targeted and environmentally friendly manner. Kelley et al. (2002) used reflectance NIRS and pyrolysis-molecular beam mass spectrometry (MBMS) analysis in conjunction with multivariate regression and PCA to differentiate brown-rot-degraded wood from non-degraded red spruce and to follow the temporal changes in wood undergoing brown-rot degradation. The red spruce samples were allowed to incubate from a week to eight months and then the wood was ground prior to spectroscopic analysis. PLSR of NIR test results versus percent weight loss for *Postia placenta*- and *Gloeophyllum trabeum*-infected spruce wood blocks yielded r of 0.96. PLSR of MBMS test results for the same samples yielded the same results. There were several changes in the NIR spectra of wood as the decay time increased. These changes included a decrease in hydroxyl vibrations associated with carbohydrates (1490 and 2100 nm) and a decrease in wood hydroxyls associated with lignin and hydrogen bonded water (1920 nm). PCA was used to differentiate noninfected wood and *P. placenta*- and *G. trabeum*-infected wood. These techniques were able to detect different types of biodegradation. (67)

The ability of tree species heartwood to resist biological degradation, referred to as "natural durability" or "decay resistance", is an important wood quality factor essential for environmentally friendly exterior timber uses. To compare the natural durability of the wood of different tree species, a classification based on the computation of "relative resistance" is highly practical. This measure is expressed in a five-class system according to the European standard EN 350-1, with classes ranging from "nondurable" (class 5) to "highly durable" (class 1). The feasibility of reflectance FT-NIRS for rapidly determining

the natural durability of the heartwood of larch trees was investigated by Gierlinger *et al.* (2003). FT-NIR spectra were collected from cross section of solid wood with fixed MC with a fiber optical probe. Basidiomycetes tests using *Coniophora puteana* and *Poria placenta* were carried out on larch heartwood, with pine sapwood (*Pinus sylvestris*) used as a reference. The relative resistance to decay based on mass loss was calculated, and durability classes were estimated according to European standard EN 350-1. PLSRs between the data sets of wood decay tests and the FT-NIR spectra were calculated. It was found that MSC of the first- or second-derivative spectra considerably improved the model predictability. High coefficients of correlation and RMSEP were obtained for cross validation based on wood decay tests with *P. placenta* ($R^2 = 0.85$, RMSEP = 0.077, range 0.27–1.13) and *C. puteana* ($R^2 = 0.97$, RMSEP = 0.078, range 0.07–1.58). Differences in the mean spectra between 1587 nm and 1887 nm were associated with the durability classes This region included first overtones from CH stretch of $=CH_2$, methyl groups, methylene groups, and aromatic substances, and OH stretch/CO stretch second overtone combinations present in all main components of wood. The flavonoid taxifolin (3,3′,4′,5,7-pentahydroxyflavanone), which is known to be a major phenolic compound in larch heartwood showed strong bands within this range. (68)

Restrictions on the use of wood preservatives have increased interest in the utilization of the natural durability of wood. Scots pine heartwood has been used in constructions exposed to risk of decay. In a laboratory experiment done by Flaete and Haartveit (2004), Scots pine wood specimens sampled from inner heartwood, outer heartwood and sapwood were exposed to the brown rot fungus *P. placenta*. Outer heartwood was found

to be more resistant than inner heartwood and sapwood. Reflectance NIR spectra obtained on a fresh radial surface of solid wood before decay testing were converted to apparent absorbance, transformed by MSC and used in multivariate calibrations to predict decay resistance of heartwood. Results from multiple test set validated PLSR models showed that resistance to decay in Scots pine heartwood could be predicted with satisfying precision using NIRS (RMSEP < 14%). (69)

Reflectance FT-NIRS was evaluated as an analytical tool for monitoring changes in milled loblolly pine samples induced by the white-rot fungus *Ceriporiopsis subvermispora* by Ferraz *et al.* (2005). Intensities of several NIR bands increased with biodegradation time, but a direct correlation between band intensities and biodegradation periods or weight losses due to the biotreatment were not observed. On the other hand, the intensities of NIR bands correlated (by PCR and PLSR) with properties reflecting the macromolecular characteristics of the components of the biotreated wood samples. Bio-kraft pulps were also characterized by wet chemical analysis and NIRS. In this case, appropriate prediction models related NIR spectral information with the chemical composition of the pulps, R^2 values ranged from 0.96 to 0.99. (70)

Weight loss, specific gravity and strength are traditional measures of how wood changes after fungal exposure. The study by Kent *et al.* (2006) investigated the effects of fungal decay (*P. placenta*) on properties of oriented strand board made of aspen including weight loss, specific gravity, dowel-bearing strength, shear strength, and alkali solubility. Predictive models using reflectance NIRS, in combination with PLSR, showed promise

as predictors of weight loss, shear strength, dowel-bearing strength, and solubility (RMSEP < 11%, R^2 > 84%). (71)

Due to their outstanding capability of degrading the recalcitrant biomacromolecule lignin, white rot fungi have been attracting interest for several technological applications in mechanical pulping and wood surface modification. However, little is known about the time course of delignification in early stages of colonisation of wood by these fungi. Fackler *et al.* (2006) used FT-NIR spectroscopic technique to monitor lignin loss of sterilised spruce wood shavings that had been degraded by various species of white rot fungi. The delignification kinetics of *Dichomitus squalens*, three *Phlebia* species, three strains of *C. subvermispora* as well as the white rot ascomycete *Hypoxylon fragiforme* and the basidiomycete *Oxyporus latemarginatus* were determined. The lignin content of spruce wood was calculated from the second-derivative of the NIR spectrum at 1673 nm which is characteristic band for CH vibrations of aromatic rings in lignin. Each of the fungi tested was able to reduce the lignin content of spruce wood significantly during the first week. The amount of delignification achieved by the selected white rot fungi after two weeks ranged from 7.2% for *C. subvermispora* to 2.5% for *P. radiata*. Delignification was significant already after 3 days of treatment with *C. subvermispora* and *P. tremellosa*. Activities of extracellular ligninolytic enzymes expressed by each of the tested fungi were determined. Lignin was degraded when peroxidase activity was detected in the fungal cultures, but only a low level of correlation between enzyme activities and the extent of delignification was found. (72)

In the study of Sykacek *et al.* (2006), reflectance FT-NIRS was investigated for its ability to predict the natural durability of commercially available larch wood. Natural durability (mass loss) tests were performed using the test brown-rot fungi *Gloeophyllum trabeum* and *C .puteana.* FT-NIR spectra were recorded on solid samples with fixed MC and transformed spectra (MSC and second-derivative) were calculated for PLSR modelling. The models were strong (R^2 = 70.7–95.19%, RMSEP = 0.063–0.026) and showed better suitability with spectra acquired from planned radial surfaces compared to those from cross section surfaces. The prediction model was valid across larch species (European and Siberian), so that exact knowledge of the growth and site conditions was not required. (73)

Fackler *et al.* (2007) cultivated beech wood veneers with white and brown rot fungi (*C. subvermispora, G. trabeum, P. placenta,* and *Trametes versicolor)* for up to 10 weeks and then fungal wood modifications were traced with FT-NIR and FT-MIR methods. PLSR models to predict the total lignin content before and after fungal decay in the range between 17.0% and 26.6% were developed for FT-MIR transmission spectra as well as for FT-NIR reflectance spectra. Weight loss of the decayed samples between 0% and 38.2% could be estimated from the wood surface using individual PLSR models for white rot and brown rot fungi, and from a model including samples subjected to both degradation types. (74)

1.3.5.2. Weathering and Aging of Wood

Tsuchikawa *et al.* (2004) used reflectance NIRS to qualitatively monitor the variation in oven-dried wood characteristics caused by UV light-irradiation. The wood species used were Japanese cypress, Japanese cedar, Sitka spruce, Japanese beech, and hackberry. It was found from the second-derivative NIR spectra that OH groups in the amorphous regions in cellulose, CH in the aromatic skeletal region due to lignin and CH in the furanose or pyranose region due to hemicellulose could be easily decomposed into low molecular mass matter by light-irradiation. The degradation rates for the amorphous regions in cellulose and hemicellulose were faster than that for lignin. Such degradation of main chemical compounds in wood by light-irradiation was closely associated with yellowing of the samples. The ratio of the second-derivative of absorbance at certain wavelengths after light-irradiation to before irradiation was investigated. In the case of OH groups in the amorphous region (1432 nm), they could not find an obvious difference in the trend between wood species. However, it was suggested that the decomposition state of lignin (1672 nm) by light-irradiation varied characteristically with wood species. Furthermore, they found a strong correlation between the ratio of absorbance due to lignin and chroma coordinates (colorfulness, brightness, and hue) measured by a colorimeter, independent of the wood species. In the case of hemicellulose, there were considerable differences in absorbance ratio (at 1712 nm) versus chroma coordinates between softwoods and hardwoods. (75)

Wang et al. (2006) reported on recent laboratory work to quantify the degradation process of wood during exposure to natural weathering. Approximately 330 southern pine lumber specimens were placed "above-ground" at an outdoor exposure site near Gulfport, Mississippi for periods up to five years. An additional 90 specimens were stored indoors to serve as controls. Two chemical preservative treatments, CCA and DDAC (Didecyldimethyl ammonium chloride), were also applied to several sets of lumber specimens. NIRS techniques on MSC transformed data were used to characterize the condition of the weathered wood surface. Multivariate statistical tools (PCA and PLSR) were used to further analyze NIR spectral data. Results indicated that NIR spectra analysis, in conjunction with multivariate statistical analysis, had good potential for monitoring changing surface conditions of wood structural members subjected to natural weathering. (76)

Yonenobu and Tsuchikawa (2003) examined the temporal changes in wood properties as estimated from NIRS applied to Japanese cypress wood from modern timber and samples collected from an old wooden building from the seventh century. The difference second-derivative spectra (difference spectra were computed by subtracting the second-derivative spectra of the antique sample from those of the modern sample) showed negative values for the absorption bands due to CH stretch first overtone in furanose or pyranose of hemicellulose at 1710 nm, and CH stretch and deformation in hemicellulose at 2330 nm. The maximum of the negative values was noticeably found at the absorption band due to OH deformation second overtone in holocellulose. These results could be directly attributed to the decrease of cellulose and hemicellulose. In contrast to

carbohydrates, the absorption band due to aromatic CH stretch first overtone in lignin at 1670 nm showed positive values in the difference spectra. The quantity of lignin may have decreased in absolute mass, but had increased relatively after 1300 years of dry exposure. The absorption bands due to OH stretch first overtone groups in the semicrystalline (1480 nm) and crystalline (1580 nm) regions in cellulose did not show significant change between the antique and the modern samples. Negative values could be observed for OH stretch first overtone groups in the amorphous region at 1430 nm. (77)

The ageing degradation of the fine wood structure of dry-exposed archaeological wood was investigated by FT-NIRS with the aid of a deuterium exchange method by Tsuchikawa *et al.* (2005). The archaeological wood sample was taken from an old wooden temple in Japan (late seventh century). Comparing the analytical results with those of a modern wood sample of the same species (Japanese cypress), the ageing process of archaeological wood was clarified as a change in the state of order on a macromolecular structural level. It was concluded from NIR spectra that the amorphous region, and partially semi-crystalline region, in cellulose, hemicellulose, and lignin decreased by the ageing degradation (1429, 1724, and 1672 nm, respectively), whereas the crystalline region in cellulose was not affected by the ageing 1550, 1591, 1790, 1830, and 2336 nm). The absorption band at 2382 nm assigned to the second overtone of the CH deformation vibration mode in holocellulose also decreased by aging. The water molecules are condensed in one or more layers on sorption sites in the amorphous region of wood. It could, therefore, be explained that the adsorption sites in wood substance

decreased by the ageing degradation (1916 nm). The accessibility of the diffusant to affect H/D-exchange was monitored by an OH-related absorption band obtained from FT-NIR transmission spectroscopy and characteristically varied with the ageing process of the wood samples, the absorption bands characteristic of a specific state of order and the diffusion agent. Finally, they proposed a morphological model to describe the variation of the fine structure of the microfibrils in the cell wall with ageing degradation. (78)

Inagaki *et al.* (2008) investigated the adsorption/desorption mechanism of water and the variation of water adsorption for modern and archaeological wood from Japanese cypress using reflectance FT-NIRS. A mixture model of water was used to decompose the NIR difference spectra into three components (free water molecules (1887 nm), those with one OH group engaged in hydrogen bonding (1923 nm), and those with two OH groups engaged in hydrogen bonding (2032 nm)) based on a PCA. The variations of each water component with relative humidity (RH) could be explained by proposing a model that described water absorption in three stages. At Stage I (RH=0–40%), water molecules interacted more strongly with wood substances than those at the other RH ranges, since monomolecular layers of water were formed. At Stage II (RH=40–90%), the water molecules might interact with adjacent water molecules, because multilayers of water molecules were formed on wood substance. At Stage III (RH=90–100%), the water molecules existed not only in an adsorbed water state but also in a bulk water state. It was shown from the NIR spectra that the mechanism of water adsorption to wood did not show a marked change due to aging. On the other hand, the modern sample adsorbed

water molecules more strongly than the archaeological wood sample. No difference in crystallinity between the modern and the archaeological samples could be found. This suggested that the aging phenomenon in wood is due to the decrease of adsorption sites in hemicellulose or amorphous cellulose. (79)

1.3.6. Pulp and Paper

1.3.6.1. Pulp Properties

A goal for the pulp and paper industry is to get a fast and reliable characterization of raw materials. One possibility for this is NIRS combined with multivariate analysis. In the study by Antti *et al.* (1996), NIRS was used to characterize a series of pine pulp samples. The pulps were characterized by 17 traditionally measured pulp properties such as porosity, tensile strength, and paper density. PLS was used for the calibration model and internal as well as external validations were done. The resulting PLS model for the 17 pulp properties gave an overall R^2 of 0.61 according to cross-validation. Predictions of the test set objects showed that most of the properties were well described by the model. (51)

The work by Marklund *et al.* (1999) presented a NIR study undertaken to explore the relationship between the choice of softwood raw material and the properties of the resulting kraft pulps expressed in terms of physical parameters for the pulps and strength properties for the corresponding handsheets. The kraft pulps were made from 20 different types of wood samples, which were chosen according to an experimental design.

Absorbance NIR spectra were recorded for the wood chips and the fully bleached pulp samples. To evaluate the relation between NIR and the end properties, PLSR analysis was used to generate prediction models for fiber properties and strength parameters. Using OSC, it was found that the NIR spectra of milled wood chips had nearly the same predictive ability as those of the bleached pulps. The PLS models for strength parameters based on NIR spectra of pulp samples used between 95% and 98% of the variation of the pre-treated NIR spectra to explain between 75% and 92% of the variation of the strength parameters using one PLS component. The predictive ability was good, corresponding to R^2 values ranging from 73.3% to 90.8%. The PLS models based on OSC treated NIR spectra of wood samples used between 77% and 90% of the variation of the OSC-corrected NIR spectra to explain between 71% and 88% of the variation of the strength parameters using one PLS component. The predictive ability corresponded to R^2 values ranging from 62.1% to 87.2%. (80)

NIRS and multivariate data analysis were applied by Fardim *et al.* (2002) to predict the chemical composition and physico–chemical characteristics of *Eucalypt grandis* unbleached kraft pulps obtained at different laboratory pulping conditions. Viscosity, degree of polymerization, kappa (it estimates the amount of chemicals required during bleaching of wood pulp to obtain a pulp with a given degree of whiteness), brightness and contents of glucan, xylan, uronic acids, and lignin were the modeled variables using reflectance NIR spectra obtained on pulp handsheets and the PLS method. Models with two to four PLS components and good predictive ability (RMSEP < 14%) were established after first-derivative spectra pre-processing and application of cross validation

methodology. The predictive models could reduce the time consuming traditional analyses in the pulping industry laboratories, and also led to a better process monitoring for suitable applications. (81)

A total of 910 maritime pine wood discs, belonging to a genetic trial of 80 families with 11–12 trees per family, were used in the study by Alves *et al.* (2007). A NIR PLSR model for the prediction of kappa number of *P. pinaster* pulps obtained from samples pulped under identical conditions was calculated. Very good correlations between NIR spectra of maritime pine pulps and kappa numbers in the range from 58 to 100 were obtained. Besides the raw spectra, spectra pre-processed with ten methods were used for PLS analysis (cross validation with 57 samples), showing that even after test set validation (with 34 samples) no model decision could be made due to almost identical statistics. The final evaluation that proved the predictive power of the models by predicting pulps with unknown kappa numbers allowed choosing a model according to a minimal number of outliers found during this process. The minimum–maximum normalized spectra in the wave number range from 1637 to 1838 nm used for the calculation gave the best model with a RMSEP of 2.3%, R^2 of 95.9%, and one PLS component. The percentage of outliers during evaluation was 0.9%. (82)

The use of calibrated Vis and NIRS to predict the Klason lignin content of Moso bamboo, Chinese fir, Paulownia, and poplar samples was investigated by Huang *et al.* (2008). For bamboo, Chinese fir, and Paulownia, the lignin content predicted by means of chemical methods and that predicted by NIR were similar. The ratio of performance to

deviation (RPD) of bamboo, Chinese fir, and Paulownia was 3.33, 2.53, and 1.77, respectively. However, for the poplar, the RPD was only 1.07. The original models were constructed using a full spectrum ranging from 400 nm to 2500 nm. If the spectral range was reduced to the range of 400–1050 nm or 1100–2500 nm, a slight decrease in the quality of the models would occur. However, this decrease was minor considering the advantages of using a reduced spectral range. The results demonstrated that NIR could predict Klason Lignin Content of bamboo, Chinese fir, Paulownia, and Poplar. (83)

1.3.6.2. Pulp Yield

Pulp yield is an important parameter for the paper industry and is function of cellulose, density, extractives, and lignin content. Each wood constituent has a different capacity to resist chemical degradation, thus influencing the variation in residual pulp per volume of wood. The perpetual goal is to increase cellulose and density and decrease the lignin content and extractives. A concept was presented by Lindgren and Edlund (1997) for monitoring the delignification process during a laboratory kraft cook on softwood (mixture of *Pinus silvestris* and *Picea abies*). Pulping of wood using the kraft process involves the cooking of wood chips in an alkaline solution at elevated temperature and pressure to dissolve lignin and leave fibers composed of cellulose and hemicellulose intact. Two series of cooking liquors were analyzed by transmittance NIRS. The first series of liquors was withdrawn from pulping experiments covering the initial, bulk and final phase of the cook. The second series of liquors was taken from experiments associated with the final delignification phase in order to be a better learning set for the

prediction of kappa numbers. MSC was applied to correct for optical interference. By using a PLSR technique, the pulp yield and the Klason lignin content could be predicted accurately from spectroscopic data obtained on the first series of liquors ($r < 97\%$). Similarly, when NIR data were used from the second series of cooking liquors, a very good correlation ($r = 98\%$) was found between observed kappa numbers and the predicted ones. (84)

Within-tree variation in kraft pulp yield, predicted using reflectance NIR analysis, was studied in 30 trees of *E. globulus* and 50 trees of *E. nitens* by Raymond *et al.* (2001) to develop a non-destructive sampling strategy. Trees, aged five to nine years, were sampled across a range of sites in southern Australia. Simulated core samples were removed at six fixed heights easily accessible from the ground and at seven percentage heights. Whole-tree values, calculated from percentage height data, were correlated with the core data to determine the optimal sampling height. The data were converted before analysis to the second-derivative mode. Core samples were found to be good predictors of whole-tree pulp yield for *E. globulus*, with simulated cores taken from the recommended sampling height (1.1 m) explaining more than 50% of variation in whole-tree pulp yield. Results for *E. nitens* were variable with large site differences apparent. On high quality sites, core samples from the recommended sampling height (0.9 m) were good predictors of whole-tree pulp yield, explaining around 60% of the variation. On poor quality sites, cores were poor predictors of whole-tree pulp yield. Radial orientation of cores was not important and predicted pulp yield was not related to tree size, basic density, or fiber length. To estimate stand mean pulp yield to an accuracy of ±1% would require sampling six trees

of *E. globulus* and four trees for *E. nitens* using either multiple discs or core samples. A single sampling height (1.1 m) was recommended for sampling for basic density, fiber length. Fiber coarseness, and predicted pulp yield in *E. globulus*. For *E. nitens* the recommended sampling height for basic density and fiber length was 0.7 m and 0.9 m was recommended for predicted pulp yield on good quality sites. (85)

Schimleck *et al.* (2006) investigated the addition of samples from a new site (Gog) to the 126 samples of Tasmania-wide *E. nitens* set, with the aim of improving predictive accuracy of pulp yield using reflectance NIRS. It was demonstrated that the addition of a single Gog sample to the Tasmania-wide calibration set was sufficient to greatly reduce predictive errors and that the inclusion of at least three Gog samples in the Tasmania-wide set was sufficient to give relatively stable predictive errors. The addition of different sets of five Gog samples to the Tasmania-wide calibration, however, caused predictive errors to vary between sets. The standard deviation of pulp yield for the prediction set (20 Gog samples) was important, with sets having the largest standard deviations giving the best predictive statistics. Finally, the Tasmania wide *E. nitens* calibration was enhanced using samples from a different species (*E. globulus*) and applied successfully to other *E. globulus* samples. (86)

1.4. Proposed approach and objectives

Considering all of the potential interactions of variables that can affect NIR spectra, emphasis will be placed on the investigation of relationships between NIR

characteristics of different species (spruce, fir, and pine) and the different variables that can affect practical application of this approach, e.g., wood moisture content and wood weathering.

2. Materials and Methods

2.1. Wood Samples

Finger-jointed SPF 1x2 and 2x4 lumber was purchased from both Home Depot and Rona in Canada. The lumber was then cut at finger jointed ends and the segments were identified according to eastern species (jack pine (*Pinus banksiana*), spruce (white spruce (*Picea glauca*) or black spruce (*Picea mariana*)), and balsam fir (*Abies balsamea*)) by a microscope using a key identification list. After classification, the wood pieces were cut to smaller non-uniform pieces (ca. 16 mm x 27 mm x 35 mm) and sanded to obtain smooth and uniform surfaces. All the wood samples were then air dried and stored indoors under ambient conditions.

2.2. Instrument Reproducibility

The reproducibility of the NIR spectrometer was examined by measuring the spectra of the same 20 wood samples (10 of each fir and pine samples) twice within two months. These wood samples were kept at fixed MC of 11% in a dark room to avoid effects of MC and light exposure.

2.3. Analysis of Effects of Moisture Content and High Temperature Drying

The wood samples were dried at 105°C until no further weight loss was observed. The final oven-dry weight was recorded. The wood samples were then conditioned to various MC values (air-dry to FSP) by using a controlled relative humidity chamber (Blue M Electric Company, Blue Island, Illinois, USA) where chamber air is humidified or de-humidified using an integral water reservoir. By controlling the temperature of the water in the reservoir, moisture is added or removed from the chamber air. Since the MC of wood is directly related to the relative humidity and temperature of the surrounding air, the MC of wood was altered as the temperature and RH of the chamber was varied. MC was calculated as per cent weight gain relative to oven-dry weight,

$$MC\,(\%) = \frac{(\text{current wood weight - oven - dry wood weight})}{\text{oven - dry wood weight}} \times 100 \qquad [7]$$

To evaluate the effect of drying, wood samples were dried at 80°C overnight. SPF is usually dried at high temperature, with relatively little humidity. These conditions could not be achieved in our kiln so an oven at 80°C and no humidity control was used as a relatively close approximation of the former conditions. The cross-section of conditioned wood samples was immediately measured by NIR.

2.4. Wood Weathering Analysis

The wood samples were tested for accelerated weathering exposure using an ATLAS Es25 Weather-Ometer. The samples were continuously exposed to UV irradiation (12000w xenon arc) and sprayed with deionized water (every two hours for 18 minutes) to different periods up to eight weeks. The temperature was kept constant at 25°C. Every two weeks the samples were taken out of the Weather-Ometer, air dried for 24 hours, and then their cross-section was measured by NIR. The cross-section surface of wood samples exposed to six weeks was cut (ca. 10 mm) and the newly fresh surface was sanded and measured by NIR in order to compare with the control (0 weeks) wood samples.

2.5. NIR Analysis

NIR reflection spectra were measured in a scanning spectrophotometer (Ocean Optics) at 6.8-nm intervals over the wavelength range 850–2600 nm. Fifty scans were collected and averaged into a single average spectrum. A box-car width of three was chosen. It specifies number of adjacent data points that are used to average across the spectral data. Increase in box-car width improves smoothing as it results in higher signal-to-noise but if it is too high there is a loss in spectral resolution. Spectra below 1100 nm and above 2300 nm were discarded because of high levels of noise observed due to instrumentation. The instrument white reflectance standard was spectralon (fluoro polymer resin). A fiber optic probe oriented at a right angle and 2 mm above the cross-section of wood surface was

used to illuminate the wood sample and to collect the reflectance spectra. NIR spectra were collected on one of the cross-sectional faces of each of the heartwood specimens.

This fiber optic probe consisted of a single input fiber encircled by 6 source fibers providing 360° of NIR illumination of a spot 6 mm in diameter. This arrangement ensured parallel orientation of the fibers. The outer six illumination fibers were connected to the tungsten halogen continuous light source (HL-2000, Ocean Optics).

The reflected light was collected by the centrally located input fiber which was coupled to a spectrometer. The spectrometer featured cooled InGaAs linear photodiode array detectors and acquired data as fast as 10 milliseconds. It is internally cooled for optimum signal-to-noise and sensitivity. InGaAs is a semiconductor composed of an alloy of gallium arsenide and indium arsenide. It consists of array of diodes (256 elements in this study) which act individually as a light-to-charge converter and storage apparatus. Each diode on the array corresponds to and detects a wavelength or set of wavelengths and thus the instrument is able to detect the whole NIR spectrum simultaneously. The computer then acquired and stored reflection spectra using SpectraSuite software. The dispersive optic was of the common Czerny-Turner design. NIR instrumentation is described in more detail in reference (1).

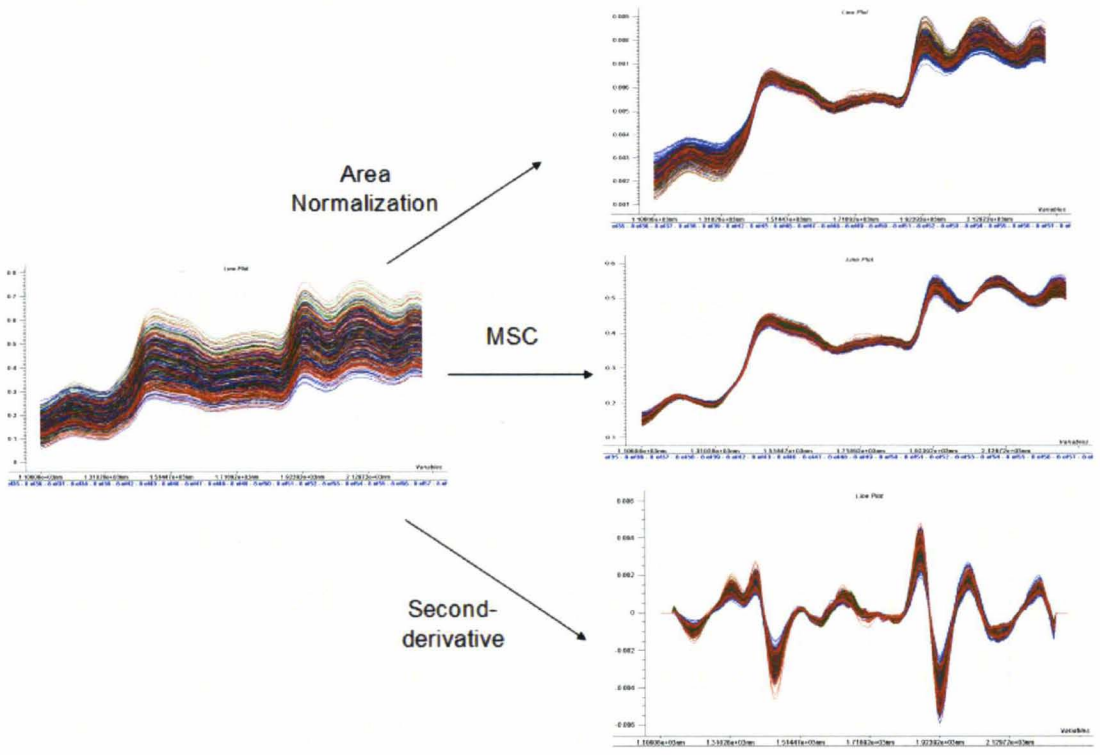
2.6. Spectral Processing

Before any spectral pre-treatment the reflection spectra was converted to absorbance and the following analysis was done on the absorbance spectra. In order to establish the best validated calibration model, several pre-processing methods were investigated. The pre-processing techniques employed were MSC, first- and second-derivative. Scatter effects were removed from the data by MSC. The hidden information in the spectra was detected by first- and second-derivatization. Spectra were smoothed and derived according to the Savitzky and Golay (1964) algorithm by means of a 7-point smoothing filter and a second-order polynomial for first-derivative and a 13-point smoothing filter and a third-order polynomial for second-derivative. Some of the pre-processing techniques were also used in combination, giving rise to a total of 6 pre-processing methods: none, first-derivative, second-derivative, MSC, MSC+ first-derivative, and MSC+ second-derivative. The best calibration model was selected by means of R^2 and rank (number of PCs). In addition, both the X- and Y-matrices were mean centered and normalised prior to performing the PLS analysis. As a rule, the first stage in multivariate modeling using projection methods is to subtract the average from each variable. This operation, called mean-centering, ensures that all results will be interpretable in terms of variation around the mean. **Fig. 2.1** shows the transformed raw spectra by various pre-processing techniques used in this study.

Variation between individual NIR spectra is the result of three main sources: non-specific scatter of radiation at the surface, variable spectral path length through the

sample, and chemical composition of the sample. In calibration one is interested only in the last source of variance. One of the main reasons for carrying out pre-processing of such data is to eliminate or minimise the effects of the other two sources. It is also possible to try to increase the differences in the contribution of each component to the total signal and in this way to make certain wavelengths more selective. The type of pre-processing depends on the nature of the signal. For this purpose, several approaches are possible.

Fig. 2.1. Raw spectra transformed by various pre-processing techniques



Normalization refers to the division of multiple sets of data by a common variable in order to eliminate that variable's effect on the data. This allows data on different scales to be compared, by bringing them to a common scale. In this study, area normalization was used. This transformation normalizes a spectrum x_i by calculating the area under the spectral curve.

$$\text{new } x_i = \frac{x_i}{\sum_j x_{ij}} \qquad [8]$$

It attempts to correct the spectra for indeterminate path length when there is no way of measuring it, or isolating a band of a constant constituent. The area under the curve becomes the same for all samples.

Multiplicative Scatter Correction (MSC) was used to transform the data in order to compensate for multiplicative and additive scatter effects in the spectra. Scatter effects are effects caused by physical phenomena. They interfere with the relationship between chemical properties and shape of a spectrum. MSC tends to simplify the calibrated model by reducing the number of PCs required, and in many cases predictions improve. The light scattering or change in path length for each sample is estimated relative to that of an ideal sample. In principle this estimation should be done on a part of the spectrum which does not contain chemical information, i.e. influenced only by the light scattering. However the areas in the spectrum that hold no chemical information often contain the spectral background where the signal-to-noise ratio may be poor. In practice the whole spectrum is sometimes used. This can be done provided that chemical differences between the samples are small. Each spectrum is then corrected so that all samples appear

to have the same scatter level as the ideal. As an estimate of the ideal sample, the average of the calibration set can be used. For each sample:

$$x_i = a + b\bar{x}_j + e \qquad [9]$$

where x_i is the NIR spectrum of the sample, and $\bar{x}_j$ symbolises the spectrum of the ideal sample (the mean spectrum) and e is the error. For each sample, a and b are estimated by ordinary least-squares regression of spectrum x_i vs. spectrum $\bar{x}_j$ over the available wavelengths. Coefficient a is the intercept (offset) of the regression line, coefficient b is the slope. Each value x_{ij} of the corrected spectrum x_i (MSC) is calculated as:

$$x_{ij}(\text{MSC}) = \frac{x_{ij} - a}{b}; j = 1,2,\ldots,\text{p.} \qquad [10]$$

Other types of signal processing used were smoothing and differentiation. By smoothing one tries to reduce the random noise in the instrumental signal. The most used chemometric methodology is the one proposed by Savitzky and Golay. It is a weighted moving window averaging method. The principle of the method is that, for small wavelength intervals, data can be fitted by a polynomial of adequate degree, and that the fitted values are a better estimate than those measured, because some noise has been removed. The user must decide the size of the window and the order of the polynomial to be used. Another way of carrying out smoothing is by repeated measurement of the spectrum, i.e. by obtaining several scans and averaging them. In this way, the signal to noise ratio, increases with $\sqrt{\text{ns}}$, ns being the number of scans. In this study, fifty scans were measured and averaged into one spectrum.

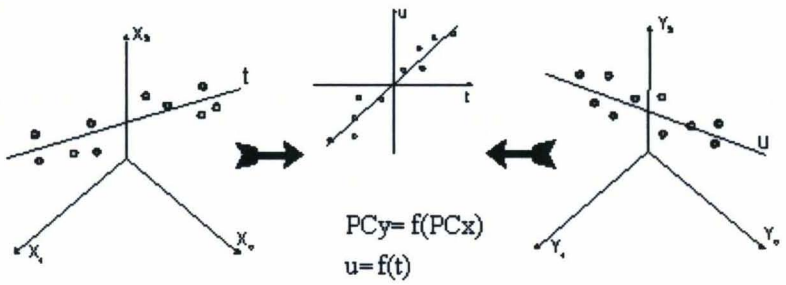
Differentiation can be used to address overlapping peaks and large baseline variations. Both first- and second-derivatives are used, but second-derivatives seem to be applied more frequently. A possible reason for their popularity is that they have troughs (inverse peaks) at the location of the original peaks. The first-derivatives removes baseline offsets while the second-derivative is used to handle both baseline offsets and scatter effects because it removes both slope and offset. In principle, differentiation of data is obtained by using the appropriate derivative of the polynomial used to fit the data in each window. One drawback of the use of derivatives is that they decrease the signal-to-noise ratio by enhancing the noise. For that reason smoothing is needed before differentiation. The higher the degree of differentiation used, the higher the degradation of the signal-to-noise ratio. Another drawback is that calibration models obtained with spectra pre-treated by differentiation are sometimes less robust to instrumental changes such as wavelength shifts which may occur over time and are less easily corrected for the changes. All spectral processing information and equations are discussed in reference (4).

2.7. Multivariate Analysis

Multivariate analysis was performed using The Unscrambler software (version 9.7, CAMO) and PLSR (Partial Least Square Regression) technique. While a complete description of multivariate analysis can be found elsewhere (4) the following summary describes the steps used to construct PLS models in this work. PLSR was applied to obtain mathematical models comparing the NIR spectra which are combined into a single data matrix (X-matrix) and the property data which are combined into a response matrix

(Y-matrix). PLSR is a bilinear modeling method where both the X- and Y-matrices are modeled simultaneously to find the latent variables (an unobserved variable that may account for variation in the data and/or for apparent relations between observed variables) in X (t_1, t_2, etc.) that will best predict the latent variables in Y (u_1, u_2, etc.). Interpretation of the relationship between X-data and Y-data is then simplified as this relationship is concentrated on the smallest possible number of components (**Fig. 4**).

Fig. 2.2. PLS Procedure (4)



The task is to describe the relationship between positions of the observations in factor space or X and their positions in response space or Y. The first component of the PLS model will orient itself so that it well describes the points in X-space while at the same time giving good correlation with the Y-space. PCs are created so that they are mutually orthogonal, thus avoiding problems related to co-linearity among the variables in the X-matrix. Compared to traditional statistical modelling based on least squares estimation, independent variables in the X-matrix are not a requirement for the PLS method. PLSR can also handle situations where the number of variables far exceeds the number of samples, which is typical for modelling with NIR data.

The Unscrambler allows for calculation of two different PLS algorithms. A PLS-1 analysis allows only one Y-variable to be projected against the X matrix at a time, while a PLS-2 analysis allows several Y-variables to be projected against the X-matrix. Both PLS-1 and PLS-2 models were constructed for different response variables.

The first stage of PLS modeling is the calibration stage where the (X, Y) relationship is established. The prediction generally is the second stage of PLS modeling. Therefore, it is mandatory to start with a known set of corresponding X and Y data. From these the relevant PLS regression model is developed. The model may subsequently be used on new X measurements to predict new Y-values. In this study, the wood samples were divided into two parts, one calibration (training) and one prediction (test) set, built up of 75% and 25% of the data respectively. For more information about multivariate analysis see reference (4).

2.7.1. Calibration of Wood Properties

During the calibration stage, a validation must be performed. Validation of a model means testing its performance according to an *a priori* given set of test result specifications. With cross validation, the same samples are used both for model estimation and testing. In this project, a full cross validation, or Leave-One-Out validation was employed. Full cross validation means that as many sub-models as there are objects are made, each time leaving out just one of the objects and only using this for testing. The squared difference between the predicted and the Y-value (residual) for each

omitted sample is summed and averaged, giving a validation Y-variance. Validation is done to determine modeling error and prediction error, outliers, optimal number of PLS components (PCs) or model dimensionality, and trends.

In order to detect sample patterns, groupings, similarities or differences, a PLS scores plot was analyzed. PLS scores are the sample coordinates along the model components. Two different sets of components can be considered, depending on whether one is interested in summarizing the variation in the X- or Y-space. The t-scores, which are the new coordinates of the data points in the X-space, are computed in such a way that they capture the part of the structure in X which is most predictive for Y; and u-scores summarize the part of the structure in Y which is explained by X along a given model component. The relationship between t- and u-scores is a summary of the relationship between X and Y along a specific model component. For diagnostic purposes, this relationship was visualized using the X-Y Relation Outliers plot. This plot visualizes the regression relation along a particular component of the PLS model. It shows the t-scores as abscissa and the u-scores as ordinate. In other words, it shows the relationship between the projection of the samples in the X-space (horizontal axis) and the projection of these samples in the Y-space (vertical axis). A sample may be outlying according to the X-variables only, or to the Y-variables only, or to both. It may also not have extreme or outlying values for either separate set of variables, but become an outlier when the (X, Y) relationship is considered. In the X-Y Relation Outlier plot, such a sample stands out as being far away from the relation defined by the other samples. If a sample stands out in such a way that it is projected far away from the center along the

model component, it is an influential outlier. Such samples are dangerous to the model: they change the orientation of the component. If there is no data transcription error for that sample, one should investigate more and decide whether it belongs to another population. If so, that sample may be removed. If not, more samples of the same kind have to be gathered, to make the data more balanced.

The number of PCs (factors) used for a model were selected by observing the response of the residual Y-variance with added factors. The residual validation variance is the mean square of all residuals, sample- or variable-wise. This is a measure of the error made when observed values are approximated by fitted values, i.e. when a sample or a variable is replaced by its projection onto the model. It expresses how much variation in the data remains to be explained once the current PC has been taken into account. The more PCs that are used in a model, the smaller the residual values of prediction validation are, but only to a point, which is the optimal number of PCs. The number of factors used in the final calibration is very important. The selection of too many factors will overfit the data and give a calibration that may only be suited to the data from which it is was derived. If not enough factors are selected, variation in the data will not be sufficiently described and underfitting occurs. The optimum number of factors for a calibration can be identified when the residual validation variance reaches a minimum and starts to plateau. However, any increase in the validation variance before the number of optimal PCs is reached is a bad sign, usually due to the presence of outliers, noise, non-linearities, etc.

The Predicted vs. Measured plot was used to check the quality of the regression model. In this plot the predicted Y-value from the model is plotted against the measured Y-value. If the model gives a good fit, the plot will show points close to a straight line through the origin and with slope equal to one. The coefficient of determination (R^2) of a regression model is a measure of the quality of the model, computed as the square of the correlation coefficient between predicted and measured values. Its value is always between 0 and 1. The closer the value to one, the better the model is. The R^2 is displayed among the plot statistics of a Predicted vs. Measured plot. When based on the calibration samples, it tells about the quality of the fit. When computed from the validation samples it tells about the predictive ability of the model. Correlation is a unitless measure of the amount of linear relationship between two variables. The correlation is computed as the covariance between the two variables divided by the square root of the product of their variances:

$$\text{Correlation} = \frac{\frac{1}{n}\sum_{j=1}^{n}\left(y_{m_j} - \bar{y}_{m_j}\right)\left(y_{p_j} - \bar{y}_{p_j}\right)}{\sqrt{\text{V}(y_m)\text{x V}(y_p)}} \qquad [11]$$

$$\text{V}(y_m) = \frac{1}{n}\sum_{j=1}^{n}\left(y_{m_j} - \bar{y}_{m_j}\right)^2 \qquad [12]$$

$$\text{V}(y_p) = \frac{1}{n}\sum_{j=1}^{n}\left(y_{p_j} - \bar{y}_{p_j}\right)^2 \qquad [13]$$

where y_m is the measured Y of sample j and $\bar{y}_m$ is the average of all Y-measured values; y_p is the predicted Y of sample j and $\bar{y}_p$ is the average of all Y-predicted values; and n is the total number of samples. Correlation varies from -1 to +1. Positive correlation indicates a positive link between the two variables, i.e. when one increases, the other has

a tendency to increase too. The closer to +1, the stronger this link is. Negative correlation indicates a negative link between the two variables, i.e. when one increases, the other has a tendency to decrease. The closer to -1, the stronger this link is. Other parameters displayed among the plot statistics of a Predicted vs. Measured plot are root mean square error, standard error, and bias.

There is an error component accumulating from each stage in the whole chain of sampling, preparation, and measurement through to data analysis. The error of either calibration or prediction validation can be expressed in several forms and is typically studied for varying number of PCs. The two most common forms are the residual Y-variance (mentioned above) or the RMSE (Root Mean Square Error), expressed in original measuring units; the latter is simply the square root of the former.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(y_j - \hat{y}_j\right)^2} \qquad [14]$$

where $\hat{y}_j$ is the predicted value of sample j, y_i is the measured value of sample j, and n is the total number of samples. Besides telling the optimal model dimensionality, the error variance also tells about the model fit and prediction ability after adding one more component, then one more, etc. RMSEP (Root Mean Square Error of Prediction) and RMSEC (Root Mean Square of Calibration) are direct estimates of the prediction error and the modelling error in Y, respectively. RMSEP is defined as the square root of the average of the squared difference between predicted and measured Y-values of the validation objects. RMSEC is the corresponding measure of the model fit, calculated from the calibration objects only. The RMSEC is a measure of the model fit, i.e., how

well has the model been fitted to the training data set. However, this does not say much about how well the model will work for new X-data. If the purpose of the modeling is prediction, the error associated with predicting new samples in the future is needed. The RMSEP is an expression of the model's ability to predict new data. As is obvious from equation [14], the overall ability is best when the prediction variance or error is at its lowest.

SEP and bias are two other statistical measures closely connected to RMSEP. Bias represents the averaged difference between predicted and measured Y-values for all samples in the validation set:

$$\text{Bias} = \frac{\sum_{j=1}^{n}(\hat{y}_j - y_j)}{\text{n}} \qquad [15]$$

Bias is a commonly used measure of the accuracy of a prediction model. Bias is also used to check if there is a systematic difference between the average values of the training set and the validation set. If there is no such difference, the bias will be zero. SEP, on the other hand, expresses the precision of results, corrected for bias:

$$SEP = \sqrt{\frac{\sum_{j=1}^{n}(\hat{y}_j - y_j - Bias)^2}{n-1}} \qquad [16]$$

If one can reasonably expect a normal distribution of the samples included in the calculation of the bias, 95% of them will have $\hat{y}_j$-$y_j \leq 2$ x SEP and 5% of them will have $\hat{y}_j$-$y_j > 2$ x SEP. This means that the uncertainty of the bias is directly dependent on

SEP. SEP increases when the Y-values (the reference values) are inaccurate. The relationship between RMSEP, SEP, and bias is:

$$RMSEP^2 \approx SEP^2 + Bias^2 \quad [17]$$

It follows that if models are unbiased SEP=RMSEP.

Finally, the regression coefficient line plot was used to determine variables that are significant in the PLS model. In a regression model equation, regression coefficients (B) are the numerical coefficients that express the link between variation in the spectra and variation in the response:

$$Y = B_0 + B_1 \text{ x } X\text{-}variable_1 + B_2 \text{ x } X\text{-}variable_2 + \ldots B_n \text{ x } X\text{-}variable_n. \quad [18]$$

For PLS, the regression coefficients can be computed for any number of components. The regression coefficient line plot shows the regression coefficients for one particular response variable (Y), and for a model with a particular number of components. Each predictor variable (X) defines one point of the line. Predictors with a large regression coefficient play an important role in the regression model; a positive coefficient shows a positive link with the response, and a negative coefficient shows a negative link. Predictors with a small coefficient are negligible and the model can be recalculated without those variables. All information and equations above are discussed in detail in reference (4).

2.7.2. Prediction of Wood Properties

After the best spectral pre-treatment was chosen, it was applied to both the calibration and the prediction sample sets. As was mentioned, calibration models were constructed with about 75% of the samples using full cross-validation. This fully cross-validated model was then used to predict the response of the test set that contained about 25% of the samples that were not included in the original model. This conservative approach ensured that the predictive capabilities of the model are reliable. The prediction models were analyzed by Y-Predicted vs. Y-Reference plot which gave the same statistical features as for the Predicted vs. Measured plot of the calibration model. Another useful plot was Predicted with Deviation. This is a plot of predicted Y-values for all prediction samples. The predicted value is shown as a horizontal line. Boxes around the predicted value indicate the deviation, i.e. whether the prediction is reliable or not. A large deviation indicates that the sample used for prediction is not similar to the samples used to make the calibration model. This is a prediction outlier meaning that the prediction sample does not belong to the same population as the samples upon which the model is based, and they cannot be trusted. The same information can be shown in a tabular form. All information above is discussed in detail in reference (4).

2.7.2.1. Partial Least Squares Discriminant Analysis (PLS-DA)

PLS-DA was used in classification of wood. PLS-DA involves developing a conventional PLSR model, but instead of a continuous variable the response variable is a

binary class indicator variable. If there are only two classes to separate, the PLS model uses one response variable, which codes for class membership as follows: 0 for members of one class, 1 for members of the other one. The PLS-1 algorithm is then used. If there are three classes or more, PLS-2 is used. Each class is represented by an indicator variable, i.e., a binary variable with value 1 for members of that class, 0 for non-members. By building a PLS-2 model with all indicator variables as Y, one can directly predict class membership from the X-variables describing the samples. The model is interpreted by viewing Predicted vs. Measured for each class indicator Y-variable: Y-predicted > 0.5 means "member"; and Y-predicted < 0.5 means "non-member". Once the PLS-2 model has been checked and validated, a prediction can be run in order to classify a new set of samples. The prediction results can be interpreted by viewing the plot Predicted with Deviations for each class indicator Y-variable: samples with Y-predicted > 0.5 and a deviation that does not cross the 0.5 line are predicted members; samples with Y-predicted < 0.5 and a deviation that does not cross the 0.5 line are predicted non-members; and samples with a deviation that crosses the 0.5 line cannot be safely classified. In this study it was known that each wood specimen was either spruce, fir or pine. For this reason three binary variables were constructed, one for each tree species. Each wood specimen was assigned to the group for which the model had the highest estimated response. All information above is discussed in detail in reference (4).

3. Results and Discussion

3.1. Spectral Pre-processing

In order to establish the best cross validated calibration model, several pre-processing methods were investigated. The cross-sections of 57 random air-dried wood samples were first measured by NIR and then classified according to species by using PLS-DA with maximum of 10 PCs. The pre-processing techniques employed on the absorbance spectra of these samples were MSC, first- and second-derivatives. Spectra were smoothed and derived according to the Savitzky-Golay algorithm by means of a 7-point smoothing filter and a second-order polynomial for first-derivative and a 13-point smoothing filter and a third-order polynomial for second-derivative. Some of the pre-processing techniques were also used in combination, giving rise to a total of six pre-processing methods: none (only normalization), first-derivative, second-derivative, MSC, MSC+ first-derivative, and MSC+ second-derivative. **Fig. 3.1- 3.6** show the Predicted vs. Measured plots for each species and a Residual Validation Variance plots for all the species together. Each Predicted vs. Measured plot was selected for the optimal PC which was determined from the Residual Validation Variance plot. The best calibration model was selected by means of the highest R^2 for the lowest number of PCs. For all species the lowest number of PCs and highest R^2 was obtained for the spectra treated by MSC+ second-derivative (**Table 3.1**). The MSC and second-derivative probably corrected for both light scattering effects such as varying path lengths due to different densities and chemical components of the wood species and instrument effects such as baseline shift (4). These pre-processing techniques were therefore applied to all following analyses.

Fig. 3.1. PLS-DA regression model overview of normalized absorbance spectra: the upper and bottom right plots are Predicted vs. Measured plots for spruce and fir, respectively, with 8 PCS; and the upper and bottom left plots are the Predicted vs. Measured plot for pine with 8 PCs and the Residual Validation Variance plot for all species, respectively.

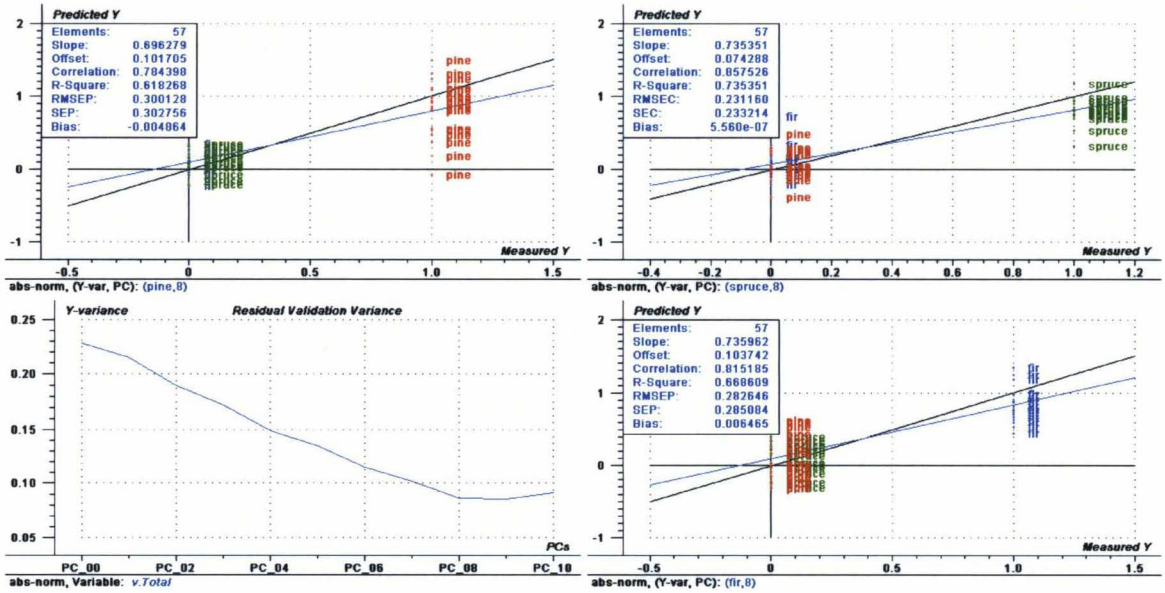


Fig. 3.2. PLS-DA regression model overview of normalized absorbance spectra transformed by MSC: the upper and bottom right plots are Predicted vs. Measured plots for spruce and fir, respectively, with 7 PCs; and the upper and bottom left plots are the Predicted vs. Measured plot for pine with 7 PCs and the Residual Validation Variance plot for all species, respectively.

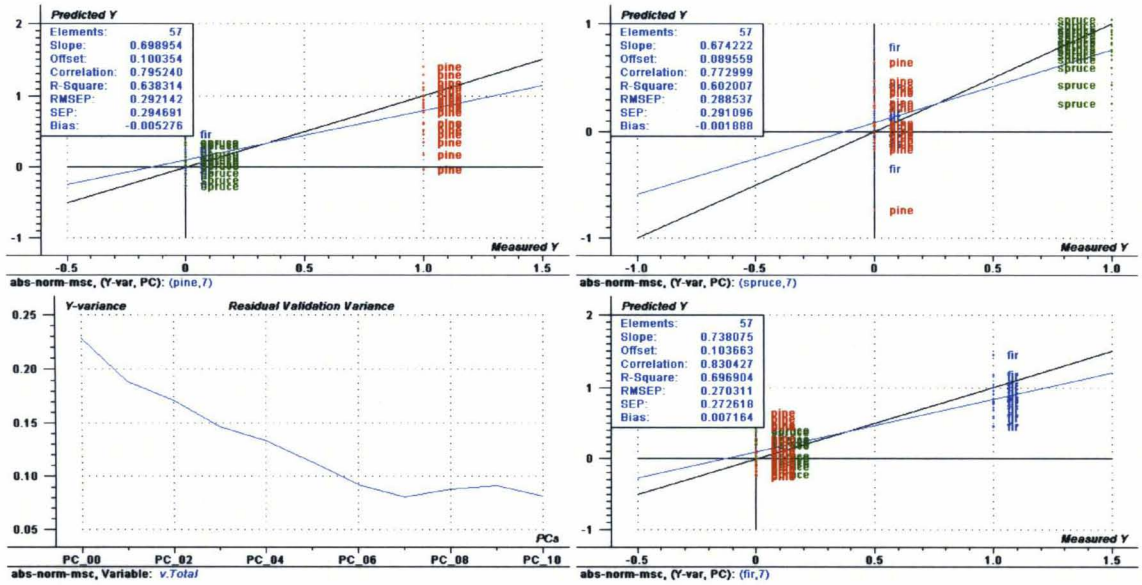


Fig. 3.3. PLS-DA regression model overview of normalized absorbance spectra transformed by first-derivative: the upper and bottom right plots are Predicted vs. Measured plots for spruce and fir, respectively, with 6 PCs; and the upper and bottom left plots are the Predicted vs. Measured plot for pine with 6 PCs and the Residual Validation Variance plot for all species, respectively.

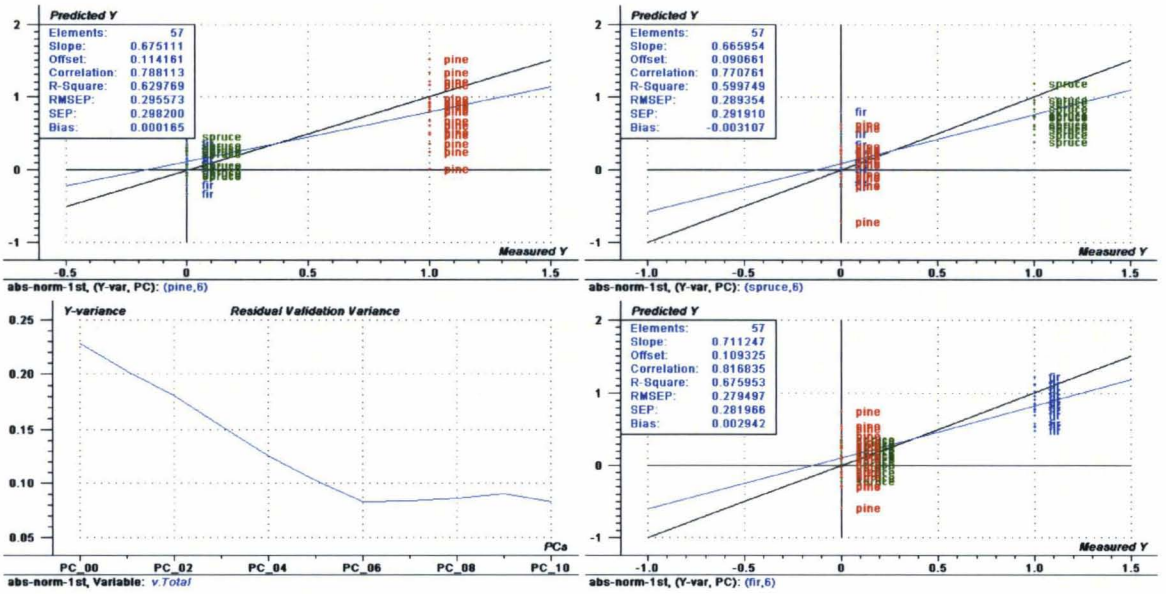


Fig. 3.4. PLS-DA regression model overview of normalized absorbance spectra transformed by second-derivative: the upper and bottom right plots are Predicted vs. Measured plots for spruce and fir, respectively, with 5 PCs; and the upper and bottom left plots are the Predicted vs. Measured plot for pine with 5 PCs and the Residual Validation Variance plot for all species, respectively.

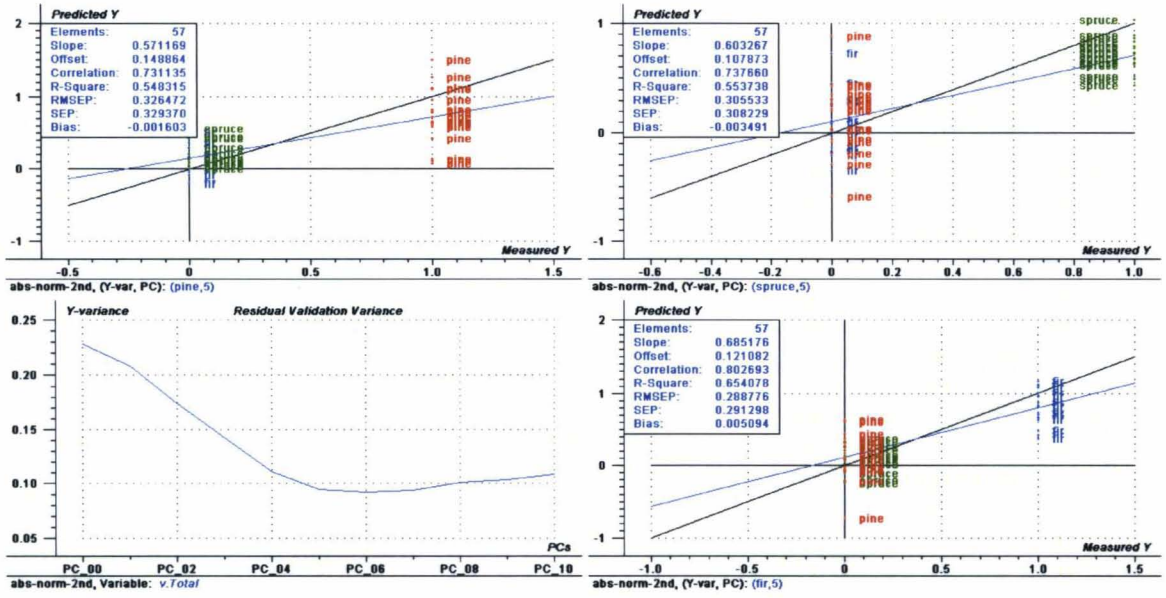


Fig. 3.5. PLS-DA regression model overview of normalized absorbance spectra transformed by both MSC and first-derivative: the upper and bottom right plots are Predicted vs. Measured plots for spruce and fir, respectively, with 7 PCs; and the upper and bottom left plots are the Predicted vs. Measured plot for pine with 7 PCs and the Residual Validation Variance plot for all species, respectively.

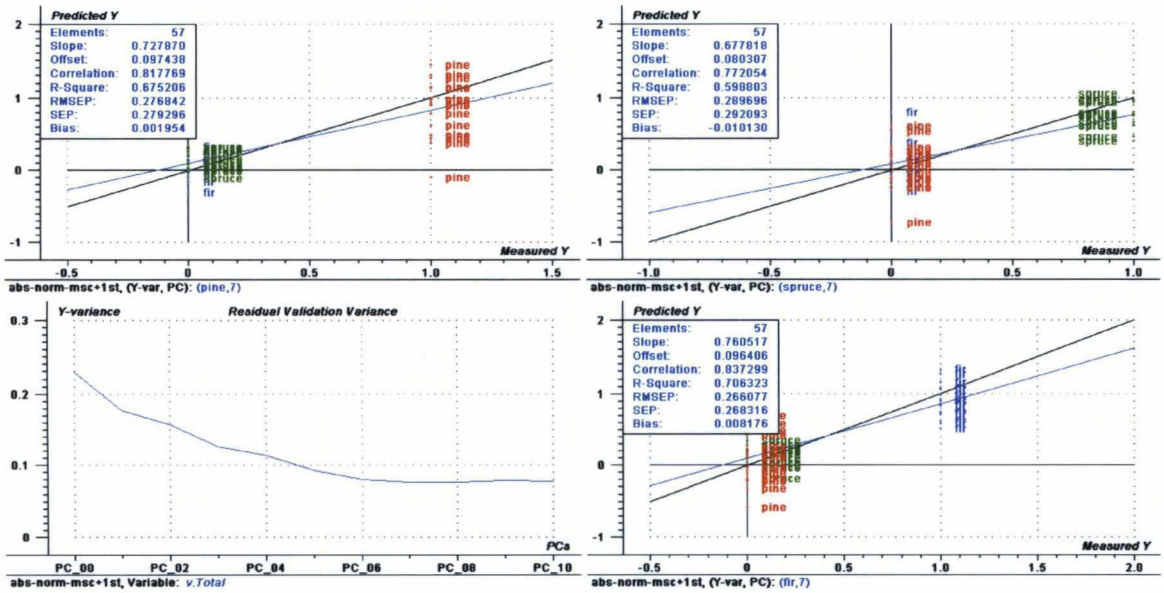


Fig. 3.6. PLS-DA regression model overview of normalized absorbance spectra transformed by both MSC and second-derivative: the upper and bottom right plots are Predicted vs. Measured plots for spruce and fir, respectively, with 5 PCs; and the upper and bottom left plots are the Predicted vs. Measured plot for pine with 5 PCs and the Residual Validation Variance plot for all species, respectively.

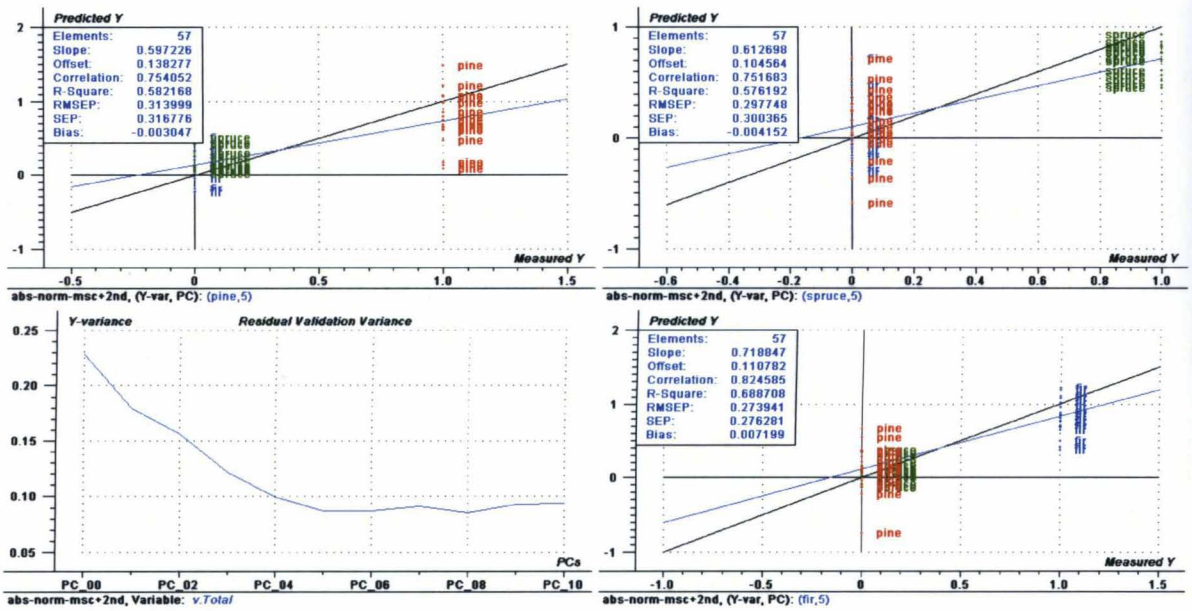


Table 3.1. The rank and the coefficient of determination values for the three wood species as determined from their PLS-DA regression models based on normalized absorbance spectra pre-treated by various techniques.

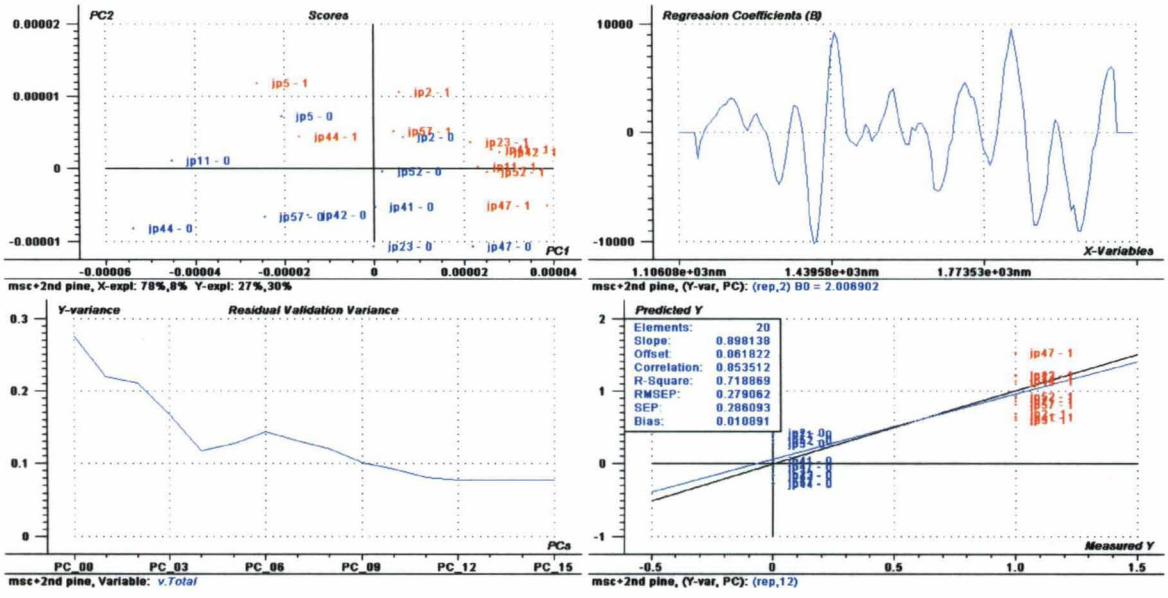
Wood Species	Spectral Processing	Rank (optimal number of PCs)	R^2 (%)
Pine	None	8	62
	MSC	7	64
	1st-derivative	6	63
	2nd-derivative	5	55
	MSC+1st-derivative	7	68
	MSC+2nd-derivative	5	58
Spruce	None	8	74
	MSC	7	60
	1st-derivative	6	60
	2nd-derivative	5	55
	MSC+1st-derivative	7	60
	MSC+2nd-derivative	5	58
Fir	None	8	67
	MSC	7	70
	1st-derivative	6	68
	2nd-derivative	5	65
	MSC+1st-derivative	7	71
	MSC+2nd-derivative	5	69

3.2. Instrument Reproducibility

The cross sections of 20 wood samples containing fir and pine kept at constant MC and inside a dark room were measured by NIR twice within two months. The samples had fixed MC to avoid the effect of moisture on the species; they were also kept inside a dark room to prevent light effects on the species. The transformed absorbance spectra of these wood samples were compared by PLS-1. **Fig. 3.7** is the validated regression overview of pine and **Fig. 3.8** of fir. As can be seen in the Residual Validation Variance plots of both of pine and fir, no separation due to NIR measurement of the same wood

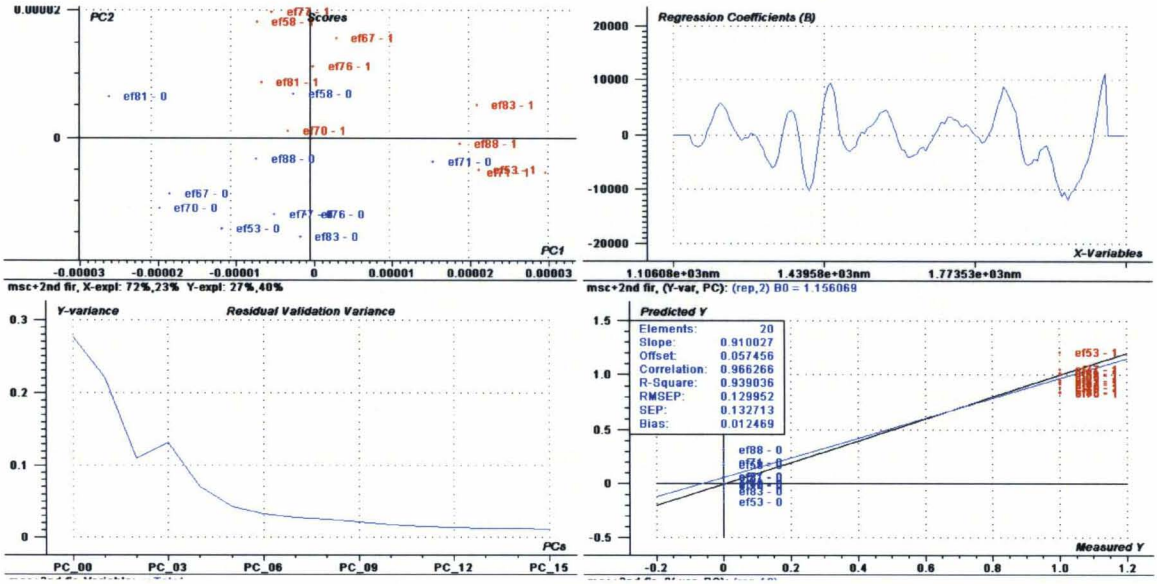
samples at different periods was possible because there was an increase in Y-variance at PC 6 for pine (**Fig. 3.7**) and at PC 3 for fir (**Fig. 3.8**). The calibration Y-variance usually decreases at all times, since the PCs are found in such a way that the residuals are minimized in each step (4). In PLS an increase in the residual variance may indicate that there is no relationship between X and Y (4). In addition, the range of the score plots for both species was very limited meaning that the samples were very similar to each other (**Fig. 3.7- 3.8**). All this lead to the conclusion that the instrument was relatively reproducible for at least two months because very close NIR absorbences were obtained for the same wood samples measured twice within two months.

Fig. 3.7. PLS-1 regression validated model of reproducibility analysis of pine samples: the upper left plot is the Scores plot; the upper right plot is the Regression Coefficients plot for PC 2; the bottom left plot is the Residual Validation Variance plot; and the bottom right plot is the Predicted vs. Measured plot with 12 PCs.



Red-pine samples measured after 2 months; blue-pine samples measured at 0 time.

Fig. 3.8. PLS-1 regression validated model of reproducibility analysis of fir samples: the upper left plot is the Scores plot; the upper right plot is the Regression Coefficients plot for PC 2; the bottom left plot is the Residual Validation Variance plot; and the bottom right plot is the Predicted vs. Measured plot with 10 PCs.



Red-pine samples measured after 2 months; blue-pine samples measured at 0 time.

3.3. Species Classification

165 air-dried wood samples (55 of each species of SPF) were conditioned to constant MC and their cross-sectional surfaces were cut and sanded (ca. 10 mm) to avoid the confounding effects of both MC and weather. They were then measured by NIRS and their absorbance spectra were treated. The spectra of these samples were divided into calibration (75%) and prediction sets (25%) and analyzed by PLS-DA in order to classify the wood samples according to species. **Fig. 3.9** shows the regression overview of the validated calibration model. The calibration model with 7 PCs was not very good having R^2 and RMSEP of 55% and 32%, 51% and 33%, and 81% and 21% for pine, spruce, and fir, respectively (**Fig. 3.9**). In this model, 57% of the spectra explained 17% of variation due to species for PC 1 which separated fir from both pine and spruce; and 12% of the

spectra explained 15% of variation due to species for PC 2 which separated spruce from pine. **Fig.3.10** and **3.11** are the regression coefficients plots for PC 1 and 2, respectively. For PC 1 the significant peaks were at 1201, 1310, 1869, 1924, 2013, 2109, 2212, and 2261 nm (**Fig. 3.10**). **Table 3.2** presents the assignment of these peaks and **Fig. 3.12** describes their trend. Fir could be separated from spruce and pine due to very small differences in the amounts and types of lignin and extractives such as starch which were higher for fir; and hemicellulose and cellulose which were higher for spruce and pine. For PC 2 the significant peaks of were at 1194, 1439, 1630, 1692, and 1931 nm (**Fig. 3.11**). **Table 3.2** presents the assignment of these peaks and **Fig. 3.13** describes their trend or lack of it. There was no clear separation between spruce and pine (**Fig. 3.13**) except for very small differentiation at 1439, 1630, and 1692 nm indicating that pine was very similar to spruce but might have different amounts and/or types of hemicellulose and/or extractives. **Table 3.3** presents the chemical composition of each of the species as was determined by various laboratories. It supports the former conclusions about species differentiation. When significant variables for species differentiation were chosen to construct a model, no improvements were seen (plots not shown) possibly because the whole spectrum was required for differentiation. In addition, no significant outliers were detected.

Since separation was not very good, calibration models containing only two species were developed and they produced better fits (**Fig. 3.14- 3.16**). These calibration models had R^2 and RMSEP of 83% and 21%, respectively for pine and fir with 4 PCs; 84% and 20% for spruce and fir with 5 PCs; and 50% and 36% for spruce and pine with 4 PCs.

These models were then used to classify wood species of a test set. Prediction results are shown in **Fig. 3.17- 3.20** and **Table 3.4**. The best prediction models were the ones that included fir and pine (R^2 = 86%, RMSEP = 19% for 4 PCs) and fir and spruce (R^2 = 74%, RMSEP = 25% for 5 PCs) where all the species were predicted correctly. The maximum error for species differentiation exhibited by the models predicting species was 12% (5 out of 42 samples were misclassified).

Fig. 3.9. PLS-DA regression validated model of all species predictions: the upper left plot is Scores plot; the upper right plot is the Predicted vs. Measured plot for pine with 7 PCs; the bottom left plot is the Predicted vs. Measured plot for spruce with 7 PCs; and the bottom right plot is the Predicted vs. Measured plot for fir with 7 PCs.

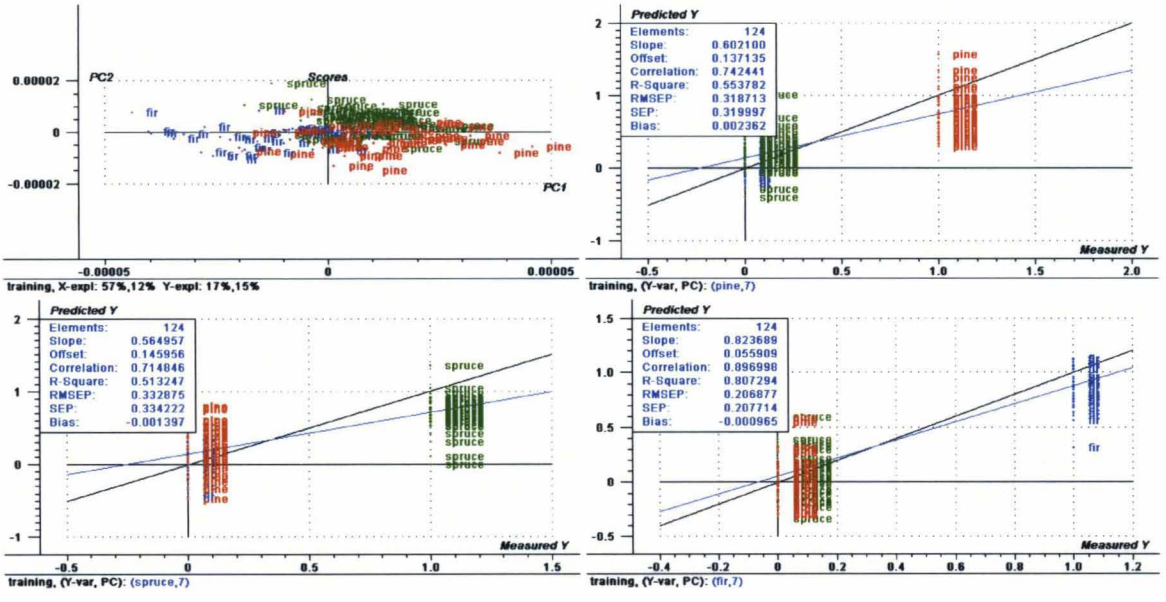


Fig. 3.10. Regression Coefficients of PC 1 from a PLS-DA regression validated model for species predictions

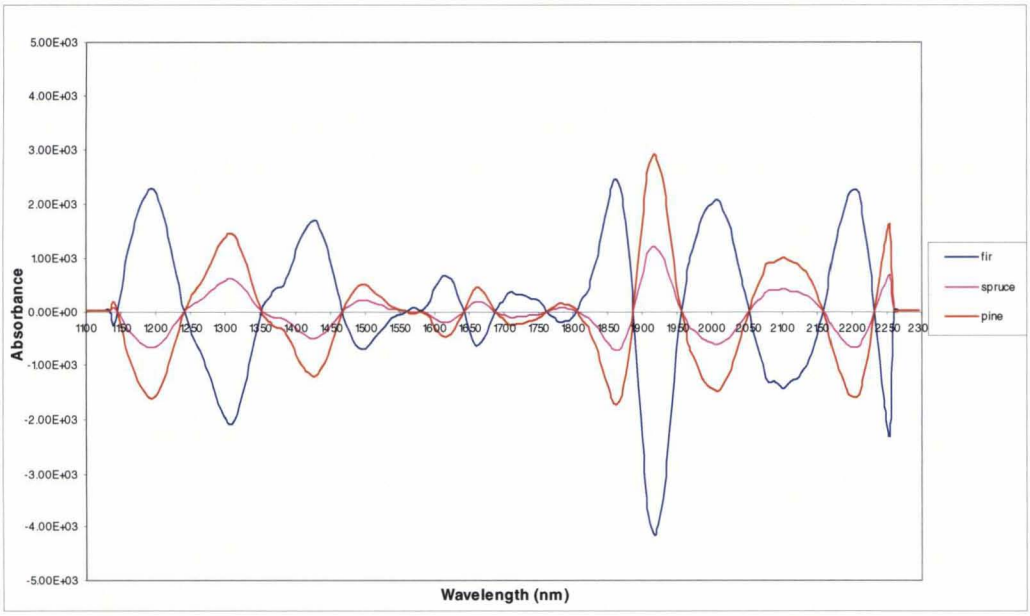


Fig. 3.11. Regression Coefficients of PC 2 from a PLS-DA regression validated model for species predictions

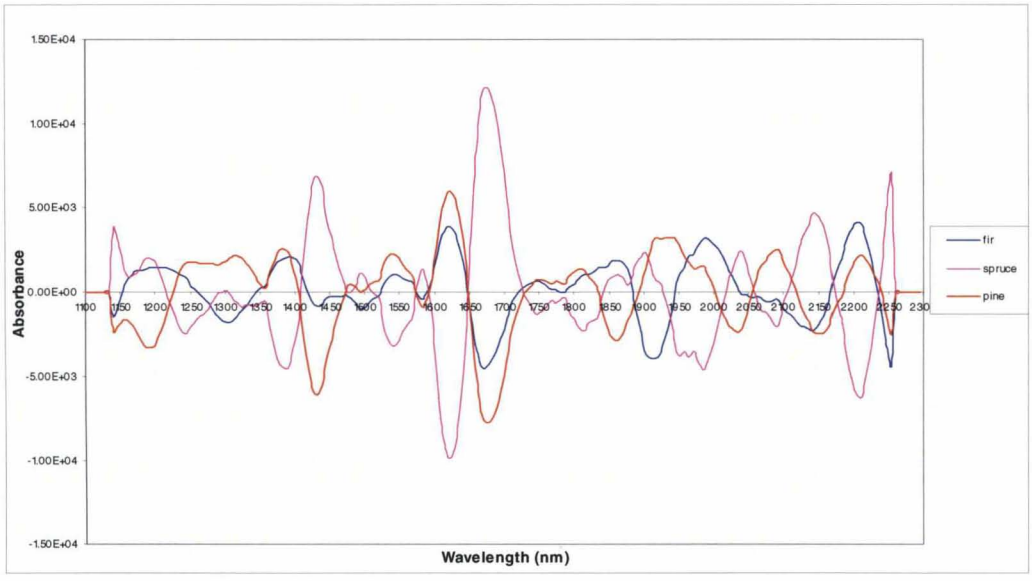
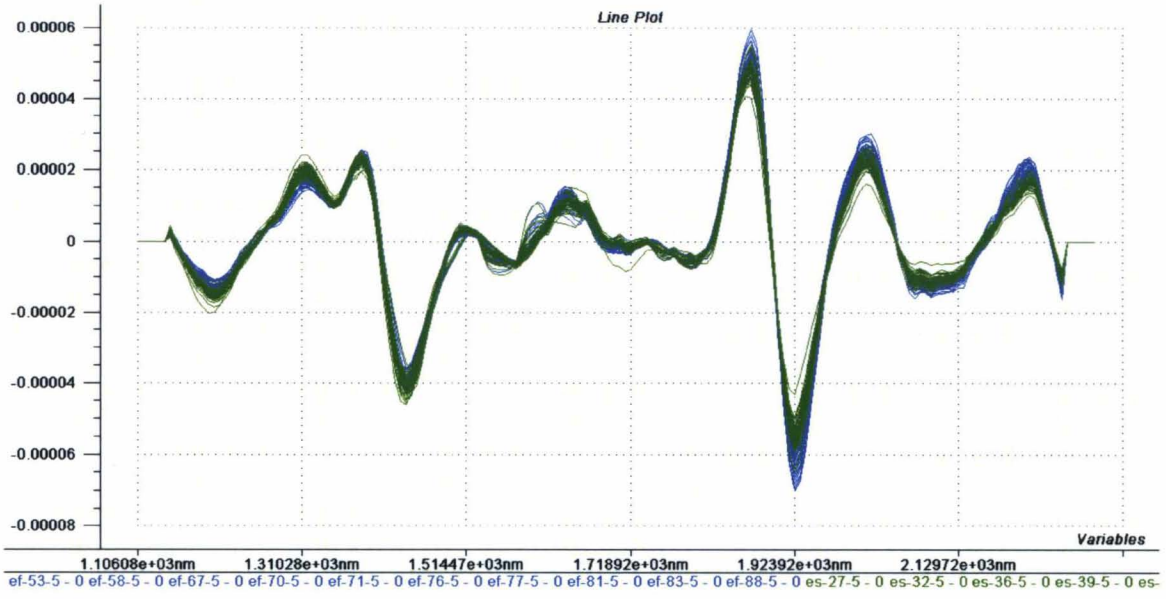
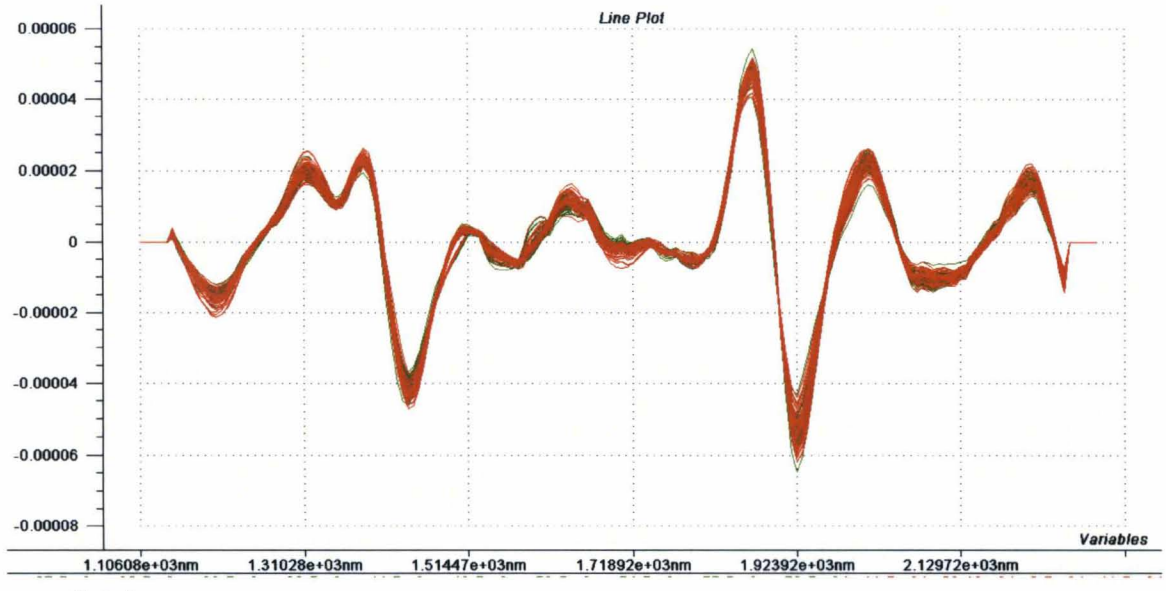


Fig. 3.12. Second derivative spectra of fir and spruce wood samples



Green- spruce; blue-fir; same plot was obtained for fir and pine.

Fig. 3.13. Second derivative spectra of spruce and pine wood samples



Red-pine; green-spruce

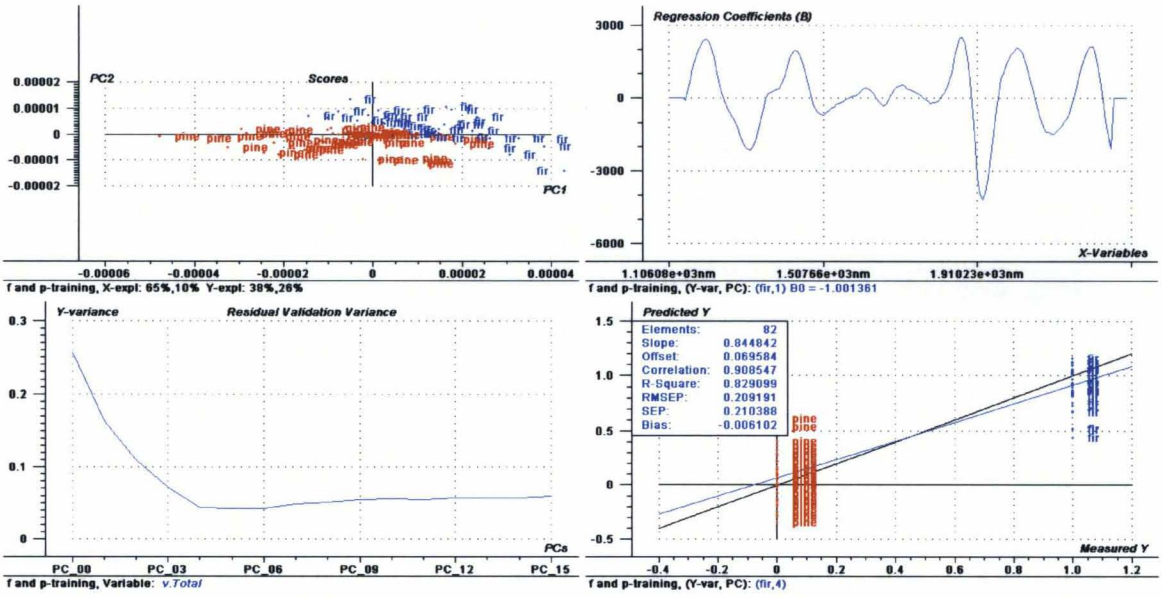
Table 3.2. Chemical assignment of the absorbance bands required for SPF separation (2).

Wavelength (nm)	Local Mode	Structure
1194, 1201	CH stretch 2nd overtone	CH_3 groups in hemicellulose
1310	CH combination	CH_3 groups in lignin
1439	CH combination	CH_2
1630	CH stretch 1st overtone	$=CH_2$
1692	CH stretch 1st overtone	Furanose/pyranose due to hemicellulose
1869	OH stretch+ CO stretch 2nd overtone combination	Cellulose
1924, 1931	OH stretch+ OH bend combination	Starch
2013	OH combination	Hydroxyl group
2109	OH bend+ CO stretch combination	Starch
2212	OH stretch+ CO stretch combination	Cellulose
2261	CH stretch+CH_2 deformation	Starch

Table 3.3. Chemical composition of eastern softwoods from the North America (percent oven-dry wood) (87)

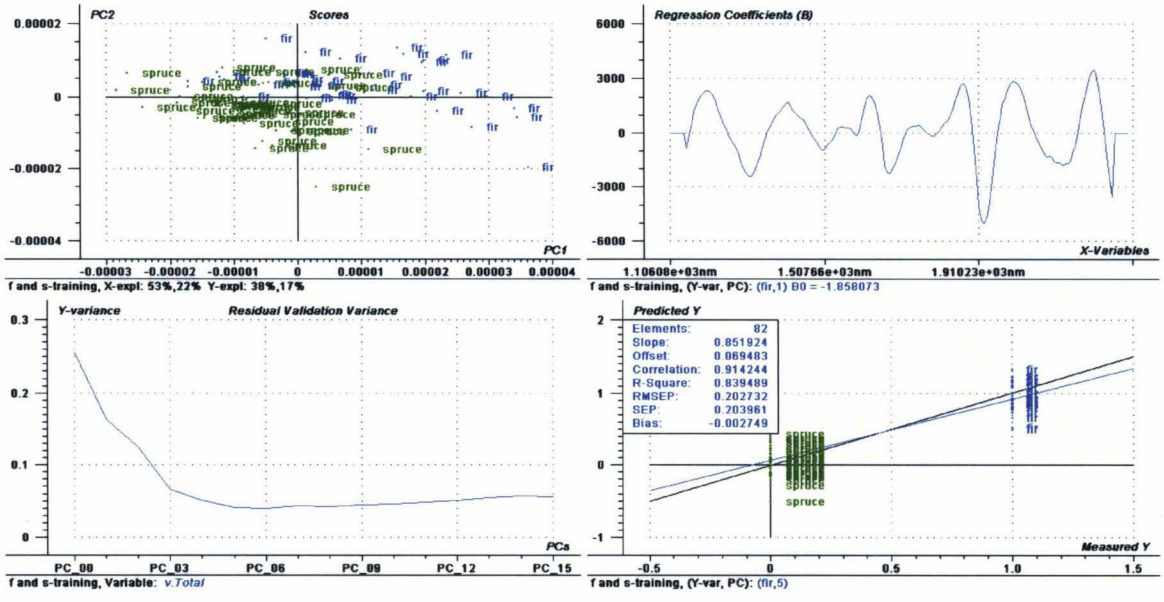
Species	Cellulose	Lignin	Ethanol/ benzene	Galactan	Mannan	Uronic anhydride	Pentosans
Balsam fir	42	29	3	1	12	3.4	11
Black spruce	43	27	2	2	9.4	5.1	12
White spruce	43	29	2	1.2	11	3.6	13
Jack pine	43	29, 27	5	1.4	10	3.9	13

Fig. 3.14. PLS-DA regression validated model of pine and fir predictions: the upper left plot is Scores plot; the upper right plot is the Regression Coefficients plot for PC 1; the bottom left plot is the Residual Validation Variance plot; and the bottom right plot is the Predicted vs. Measured plot for fir with 4 PCs.



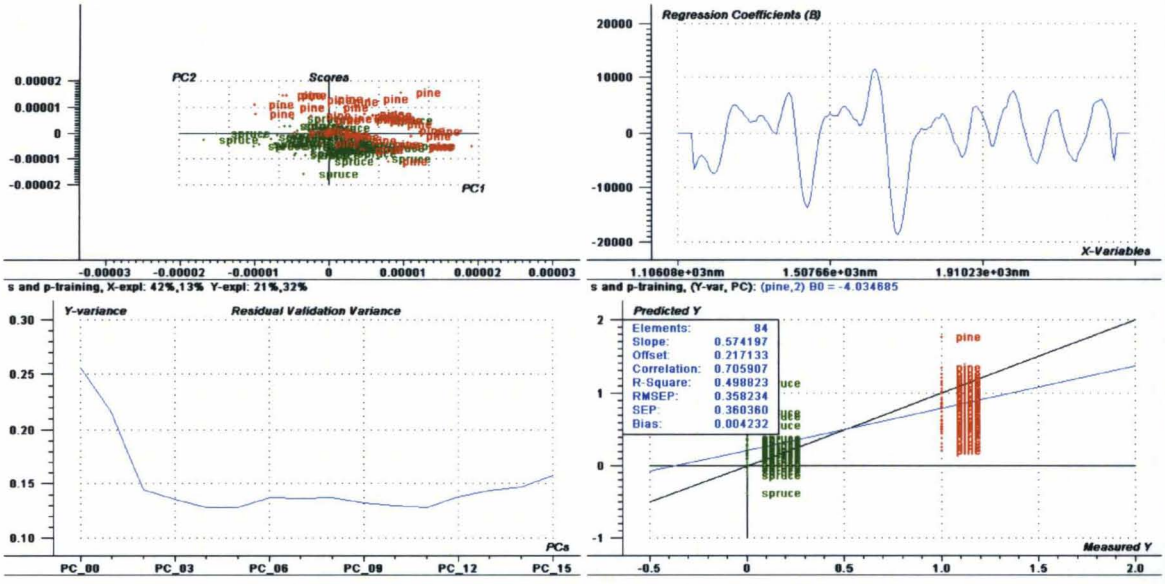
The regression coefficient plot for pine was the same but in opposite direction; and Predicted vs. Measured plot was the same for pine (not shown).

Fig. 3.15. PLS-DA regression validated model of spruce and fir predictions: the upper left plot is Scores plot; the upper right plot is the Regression Coefficients plot for PC 1; the bottom left plot is the Residual Validation Variance plot; and the bottom right plot is the Predicted vs. Measured plot for fir with 5 PCs.



The regression coefficient plot for spruce was the same but in opposite direction; and Predicted vs. Measured plot was the same for spruce (not shown).

Fig. 3.16. PLS-DA regression validated model of spruce and pine predictions: the upper left plot is Scores plot; the upper right plot is the Regression Coefficients plot for PC 2; the bottom left plot is the Residual Validation Variance plot; and the bottom right plot is the Predicted vs. Measured plot for pine with 4 PCs.



The regression coefficient plot for spruce was the same but in opposite direction; and Predicted vs. Measured plot was the same for spruce (not shown).

Fig. 3.17. PLS-DA regression prediction model of all species: the upper left plot is Predicted vs. Reference plot for fir with 7 PCs; the upper right plot is the Predicted vs. Reference plot for pine with 7 PCs; and the bottom left plot is the Predicted vs. Reference plot for spruce with 7 PCs.

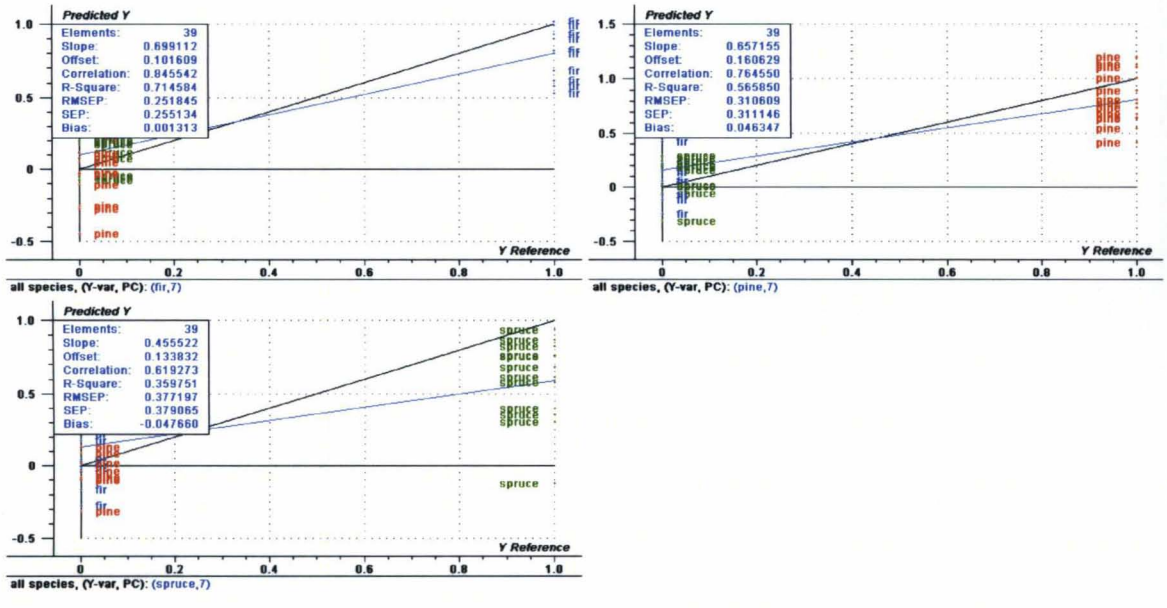


Fig. 3.18. PLS-DA regression prediction model of fir and pine: the upper plot is Predicted vs. Reference plot for pine with 4 PCs; and the bottom plot is the Predicted vs. Reference plot for fir with 4 PCs.

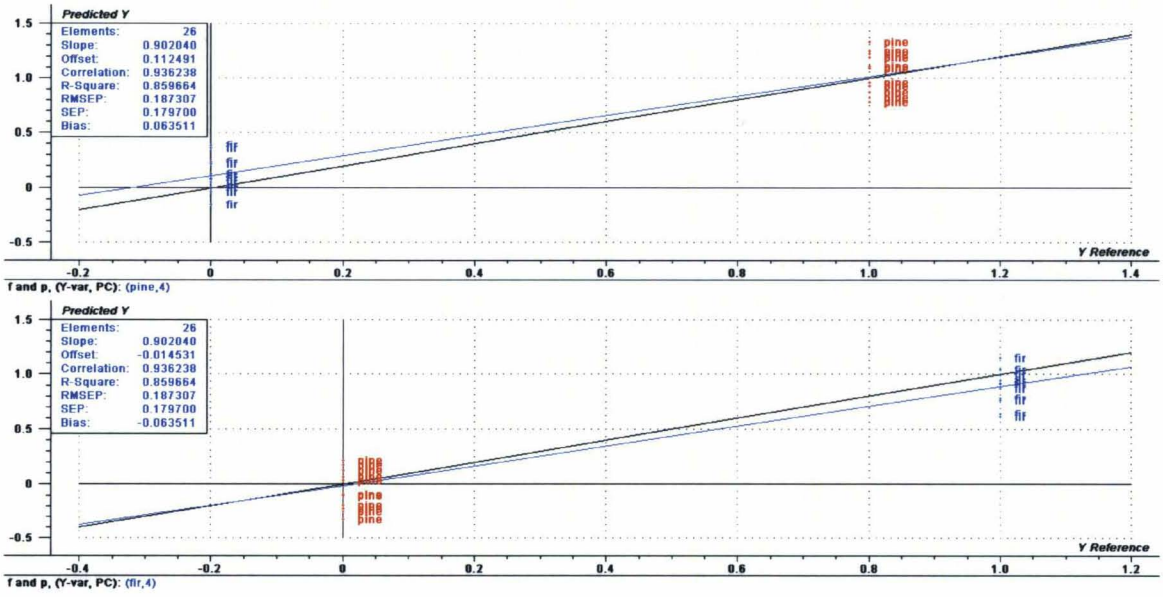


Fig. 3.19. PLS-DA regression prediction model of fir and spruce: the upper plot is Predicted vs. Reference plot for spruce with 5 PCs; and the bottom plot is the Predicted vs. Reference plot for fir with 5 PCs.

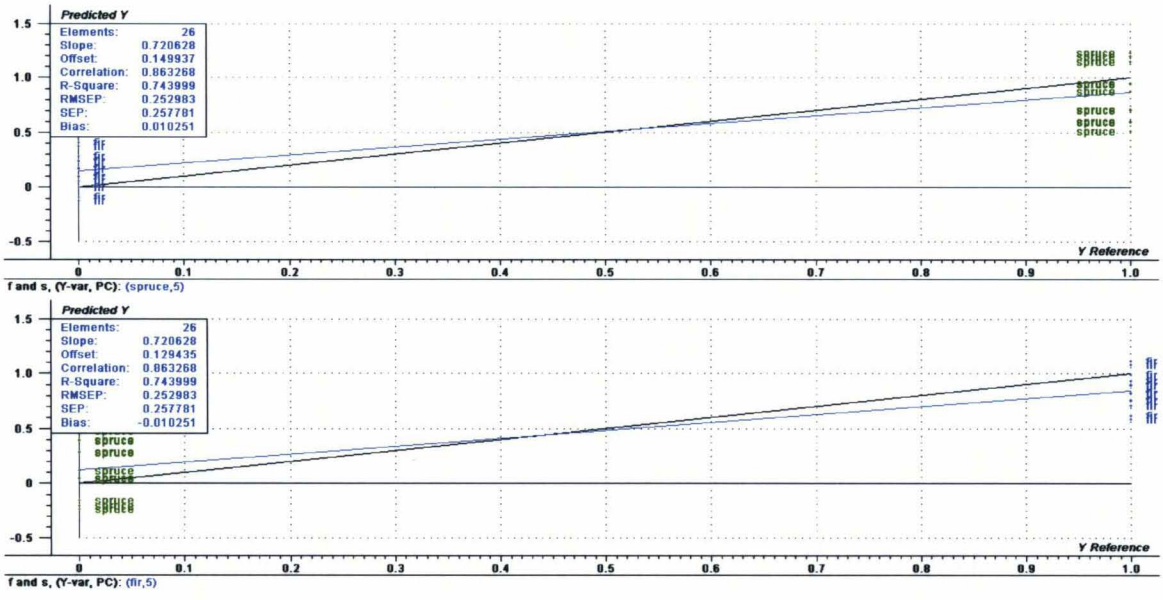


Fig. 3.20. PLS-DA regression prediction model of spruce and pine: the upper plot is Predicted vs. Reference plot for spruce with 4 PCs; and the bottom plot is the Predicted vs. Reference plot for pine with 4 PCs.

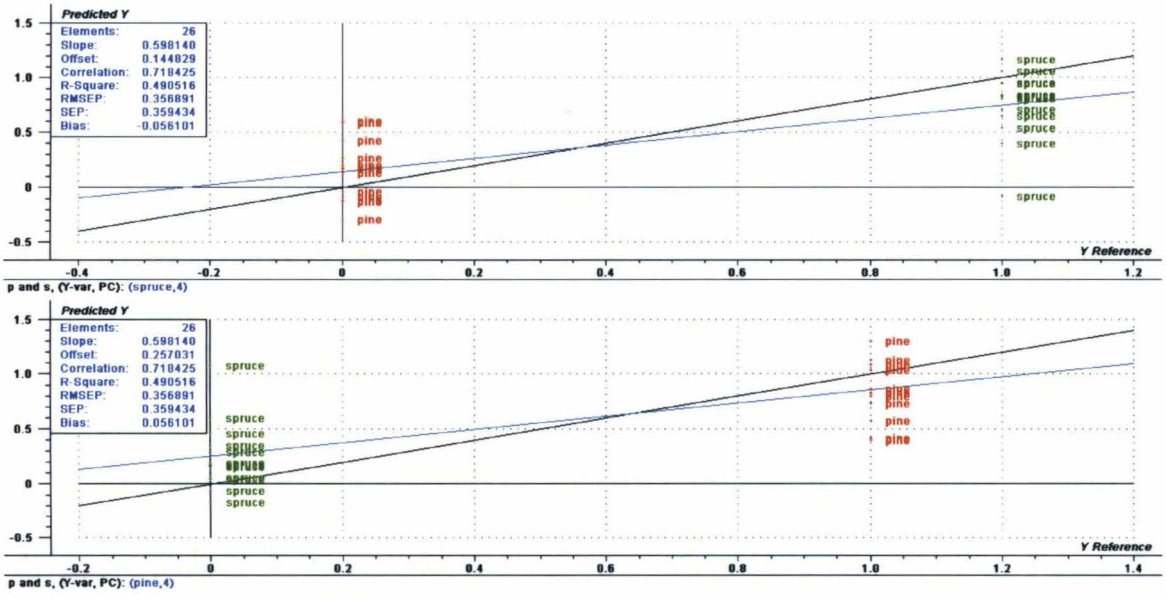


Table 3.4. Prediction results for the validated PLS-DA calibration models that classified species.

Species included in the model	# of correct species type predictions (/14)			# PCs	R^2 (%)			RMSEP (/1)		
	Pine	Fir	Spruce		Pine	Spruce	Fir	Pine	Spruce	Fir
Spruce, fir, and pine	13*	14	10*	7	56	36	71	0.31	0.38	0.25
Pine and fir	14	14	N/A	4	86	N/A	86	0.19	N/A	0.19
Fir and spruce	N/A	14	14	5	N/A	74	74	N/A	0.25	0.25
Pine and spruce	12**	N/A	12**	4	49	49	N/A	0.36	0.36	N/A

*1 pine was assigned as spruce; 3 spruces were not assigned to any of the species and 1 spruce was assigned as pine. 4 **2 pines were assigned as spruce and 2 spruces were not assigned to any of the species.

3.4. Analysis of Effects of Moisture Content and High Temperature Drying

The wood of living trees contains a large amount of water, which often constitutes more weight than the actual wood. Water has a significant influence on wood: wood continually exchanges moisture with its surroundings. Changes in response to daily humidity fluctuations are small and usually of no consequence. However, changes that occur as a result of seasonal variation, although gradual, tend to be of more concern.

Moisture in wood exists in two forms: free water, liquid filling the wood cell cavities which is held by capillary forces and bound water, liquid or vapour chemically bound (adsorbed) by hydrogen bonding to the wood cell walls. The attraction of wood for water arises from the presence of free hydroxyl (OH) groups in the cellulose, hemicellulose, and lignin molecules in the cell wall. Water molecules are first adsorbed monomolecularly (one per site) onto the wood. At higher MCs, additional water molecules first bind to the remaining sorption sites and then begin to polymolecularly adsorb onto each other. Therefore, hydroxyl groups of adjacent molecules can be mutually bonded. (6, 87)

As wood dries, the free water in the cell cavities is drawn away first. Once the free water is removed, the bound water is gradually released from the cell walls. The MC at which all of the free water is removed from the cell cavities but the cell walls are still completely saturated is called fiber saturation point (FSP). At MC above the FSP the

physical and mechanical properties of wood remain constant as MC changes. At MC below the FSP the physical and mechanical properties of wood change as MC changes. Below the FSP, the forces holding the bound water to the wood become greater as MC decreases. As wood approaches the dry condition, most moisture is adsorbed in a monolayer. As drying continues the hydroxyl groups of wood components such as cellulose move closer together forming weak cellulose-to-cellulose bonds. (5, 6, 87)

Because of the sorptive nature of wood, it has the ability to exchange water vapour with the surrounding air until it obtains moisture equilibrium with the air. Thus, wood is called a hygroscopic material. As the RH drops, it loses water; as the RH increases, the wood regains water. For a given RH level, a balance is eventually reached at which the wood is no longer gaining or losing moisture. When this balance of moisture exchange is established, the amount of water eventually contained in a piece of wood is called the equilibrium moisture content (EMC) of the wood. The EMC of wood varies directly with the ambient RH significantly and, to a lesser degree, inversely with the temperature. (5, 87)

Water in wood normally moves from high to low zones of MC, meaning that the wood surface must be drier than the interior if moisture is to be removed. Drying can be divided into two phases: movement of water from the interior to the wood surface and evaporation of water from the surface. The surface fibers of most species reach moisture equilibrium with the surrounding air soon after drying begins. This is the beginning of the development of a typical moisture gradient. (5)

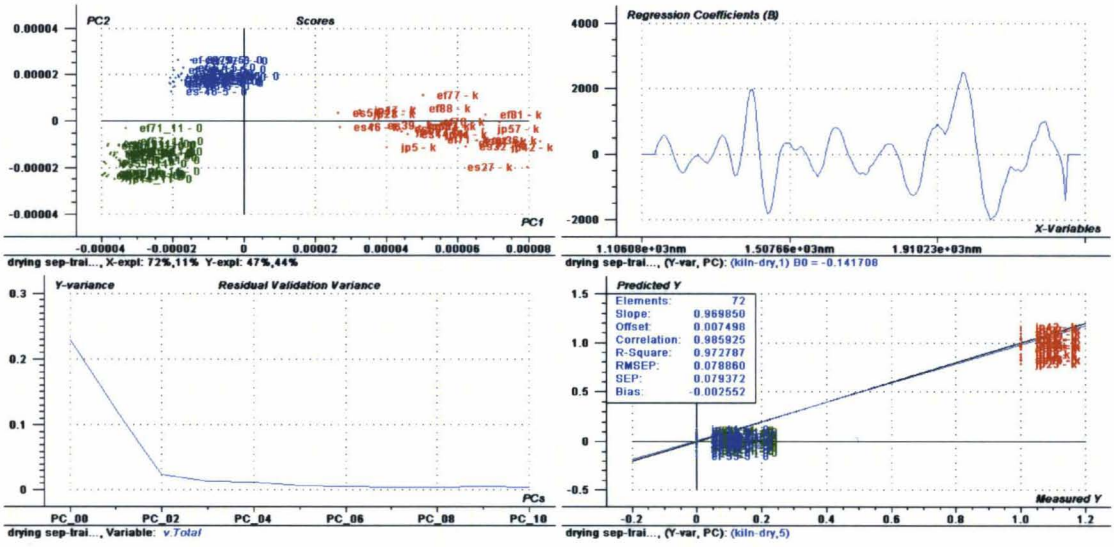
Water moves through the interior of wood as a liquid or vapour through various air passageways in the cellular structure of the wood, as well as through the wood cell walls. Water moves by two major mechanisms: capillary action (liquid) and diffusion of bound water (vapour). The rate at which moisture moves in wood depends on the RH of the surrounding air, the steepness of the moisture gradient, and the temperature of the wood. (6, 87)

A problem with NIR analysis of samples with high MC is the presence of several strong absorption bands in the NIR spectrum caused by water. These broad peaks obscure spectral information derived from some wood components. The most obvious pure water absorption peaks are the O-H first stretch overtone band at 1450 nm and the O-H stretch and O-H bending combination bands at 1930 nm. Each of these broad bands contains information on more than one hydrogen-bonded subspecies. It is known that variations in hydrogen-bonded molecular subspecies can cause band broadening and peak position shifts. Variations in hydrogen bonding (intra- and intermolecular) result in changes in the force constants of the X-H bonds. For example, when a hydrogen bond is formed, the force constant for the covalent O-H bond will decrease and the frequency of the O-H absorption band will decrease (equation [2] and [3]). The weak hydrogen bond is readily disrupted by thermal collisions and the O-H absorption band is broadened. (2)

3.4.1. Effect of Drying Analysis

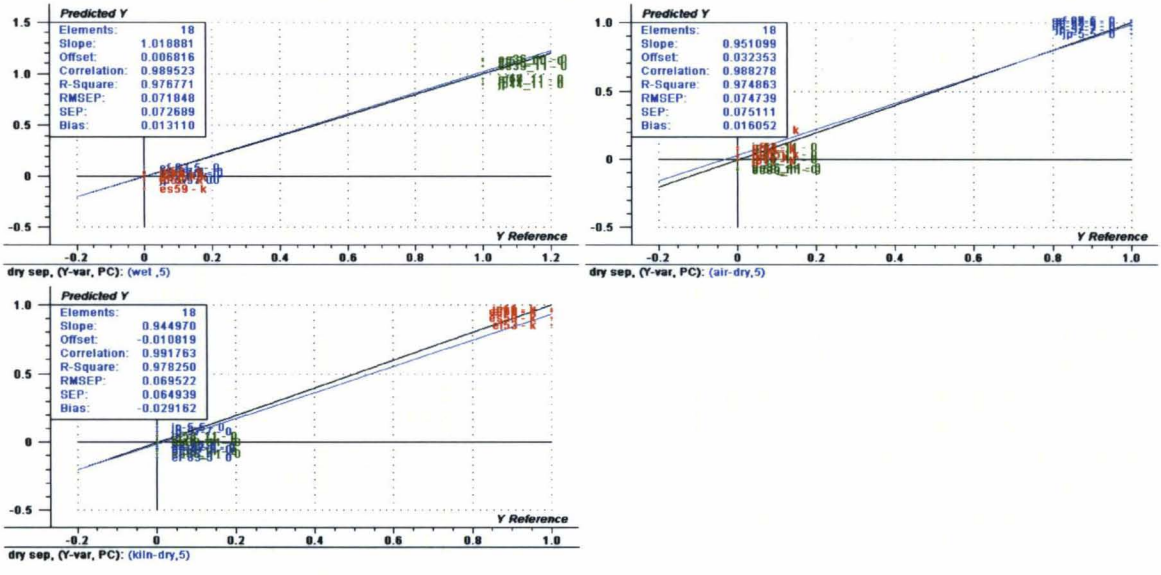
The cross sections of 30 oven-dried (MC=2%) wood samples (10 of each species of SPF) were measured by NIRS. Their transformed absorbance spectra were compared to both air-dried wood samples (MC=8%) and slightly wetter wood samples (MC=10%) by using PLS-DA. The validated calibration model overview is shown in **Fig. 3.21** and the prediction results in **Fig. 3.22.** The scores plot showed clustering of wood samples with a clear distinction between samples of high and low MCs (**Fig. 3.22**). The calibration model with 5 PCs was good having R^2 and RMSEP of 97% and 8%, respectively (**Fig. 3.21**). In this model, 72% variation in spectra explained 47% of variation due to drying condition of wood samples for the first PC (**Fig. 3.21**). The significant variables required for the separation due to drying conditions were at 1405 and 1979 nm (negative peaks); 1446, 1849, 2068 nm (positive peaks) (**Fig. 3.21**). **Fig. 3.23** shows the raw spectra of oven-dried and MC=10% wood samples. The absorption peak at 1930 nm decreased upon drying meaning that only water vibrations were involved. There was an overlap between oven-dried and 10% MC wood samples at 1460 nm meaning that other than water vibrations were involved i.e., the OH stretch first overtone of cellulose. All the samples in the test set were predicted accurately with R^2 and RMSEP of 98% and 7%, respectively for 5 PCs independently of the wood species analyzed (**Fig.3.22**). No outliers were detected. When the former significant variables were used to construct both calibration and prediction models these models got worse (plots not shown).

Fig. 3.21. PLS-DA regression validated model of drying condition predictions: the upper left plot is Scores plot; the upper right plot is the Regression Coefficients plot for PC 1 for oven-dried samples; the bottom left plot is the Residual Validation Variance plot; and the bottom right plot is the Predicted vs. Measured plot for oven-dry with 5 PCs.



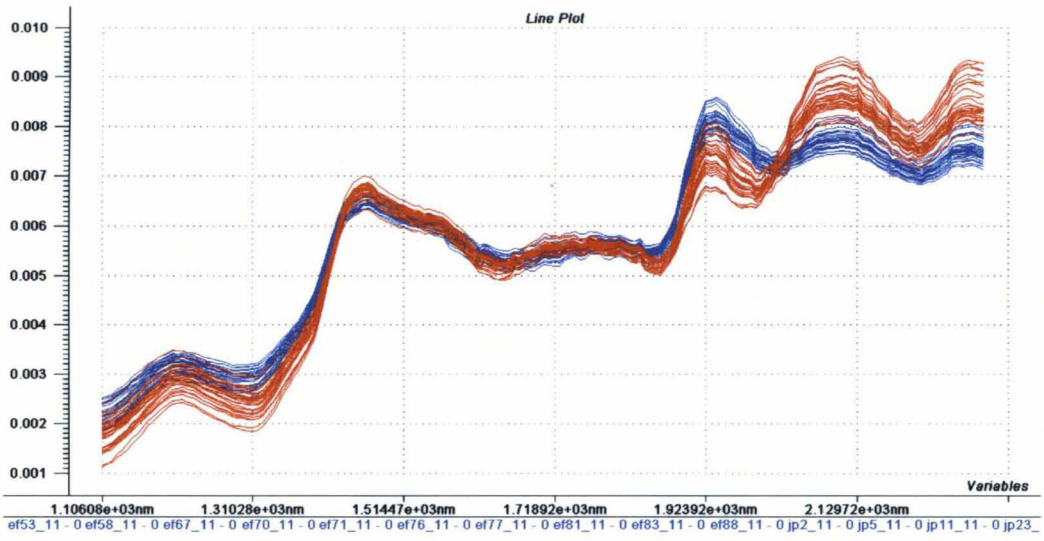
Blue- air-dried (8%) samples; green-wetter (10%) samples; and red- oven dried samples. The regression coefficient plots were the same for air-dry and wetter samples but in opposite direction; and Predicted vs. Measured plots were similar (not shown).

Fig. 3.22. PLS-DA regression prediction model of drying conditions of all species: the upper left plot is Predicted vs. Reference plot of wet wood samples with 5 PCs; the upper right plot is Predicted vs. Reference plot of air-dried wood samples with 5 PCs; and the bottom plot is the Predicted vs. Reference plot of oven-dried wood samples with 5 PCs.



Blue- air-dried samples; green-wetter samples; and red- oven dried samples

Fig. 3.23. Raw spectra of oven-dried and 10% MC wood samples



Blue- 10% MC wood samples and red- oven dried wood samples

The next step was to determine if variation due to drying conditions could affect species separation. In order to test that, PLS-DA calibration validated models were constructed based on species separation at all drying conditions. When all three species were included in the model with 11 PCs, R^2 and RMSEP were 83% and 19%, respectively for pine; 70% and 26% for spruce; and 88% and 22% for fir (**Fig. 3.24**). In the scores plot, fir and spruce were separated from pine by PC 3 and fir was separated from spruce by PC 1 (**Fig. 3.24**). In this model, only 24% variation in spectra explained 19% of variation due to species differentiation for the first PC (**Fig. 3.24**). **Fig. 3.25** shows the significant variables needed for the three species separation (discussed in section 3.3) for PC 3. Calibration models containing only two species were also constructed. They produced better models with R^2 and RMSEP of 77% and 24%, respectively with 6 PCs for spruce and fir; and 85% and 20% with 8 PCs for pine and fir

(**Fig. 3.26- 3.27**). As can be seen in Residual Validation Variance plot, no separation was possible for spruce and pine due to increase in the residual variance at PC 1 (**Fig. 3.28**). No significant outliers were detected.

Fig. 3.29- 3.31 and **Table 3.5** include the prediction results for all the models constructed. The best prediction model was obtained for fir and pine (R^2 = 85%, RMSEP = 19% with 8 PCs) where all the species were determined correctly. Despite the relatively low R^2 values and high RMSEP of models created on the training sets, the maximum error for species differentiation exhibited by the test sets was 7% (2 out of 27 samples were misclassified).

Fig. 3.24. PLS-DA regression validated model of all species predictions subjected to various drying conditions: the upper left plot is Scores plot; the upper right plot is the Predicted vs. Measured plot for spruce with 11 PCs; the bottom left plot is the Predicted vs. Measured plot for pine with 11 PCs; and the bottom right plot is the Predicted vs. Measured plot for fir with 11 PCs.

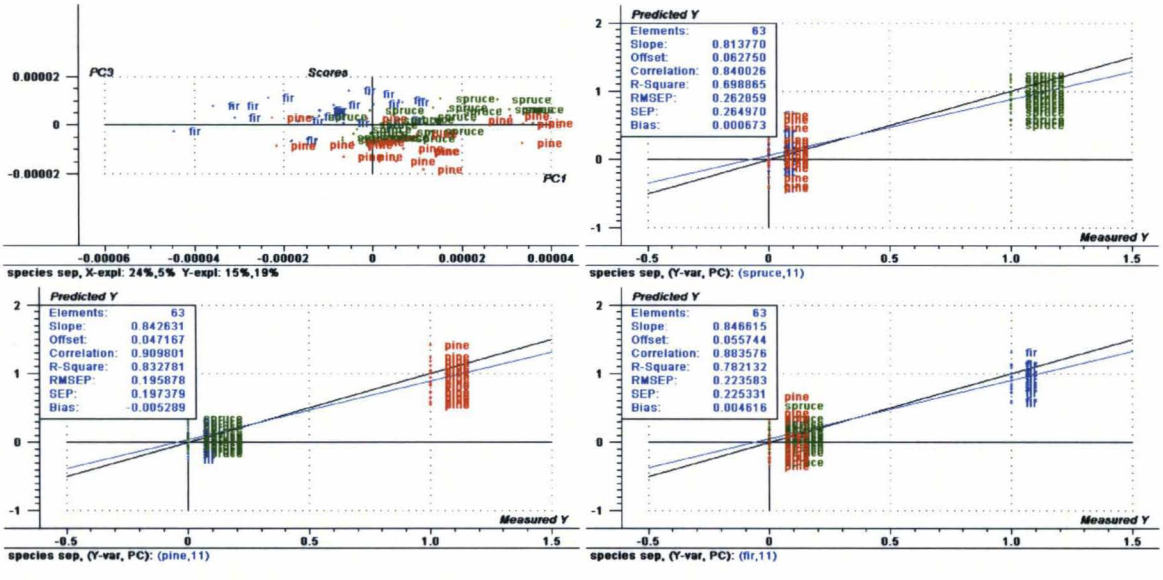


Fig. 3.25. Regression Coefficients of PC 3 from a PLS-DA regression validated model for species predictions subjected to various drying conditions

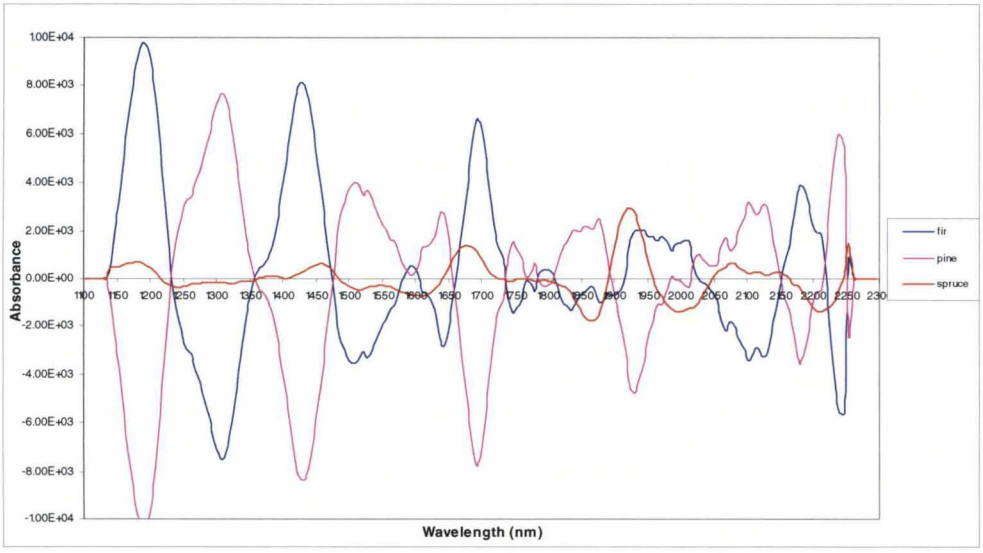
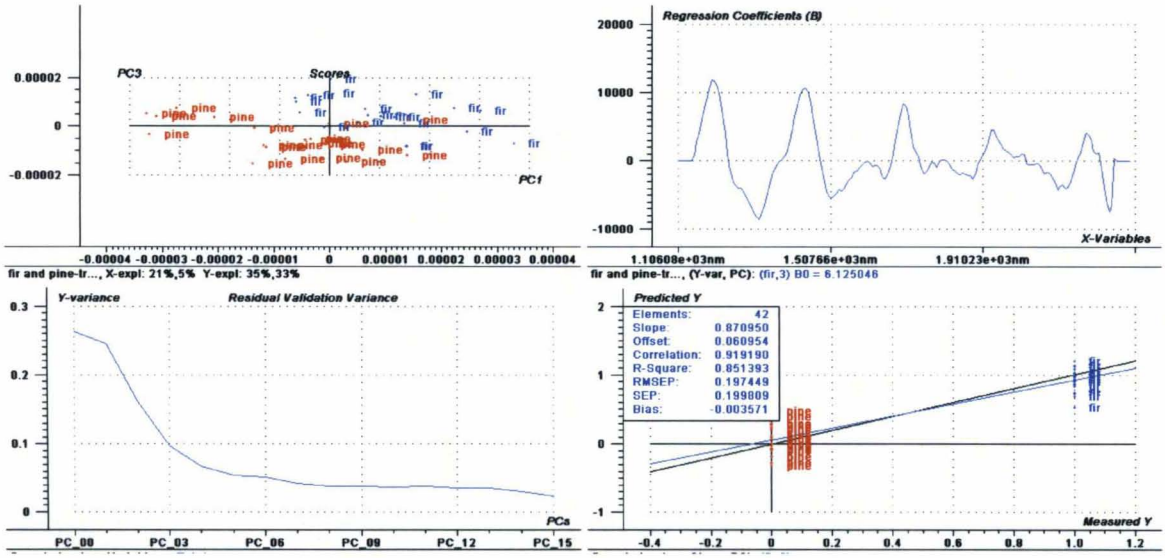
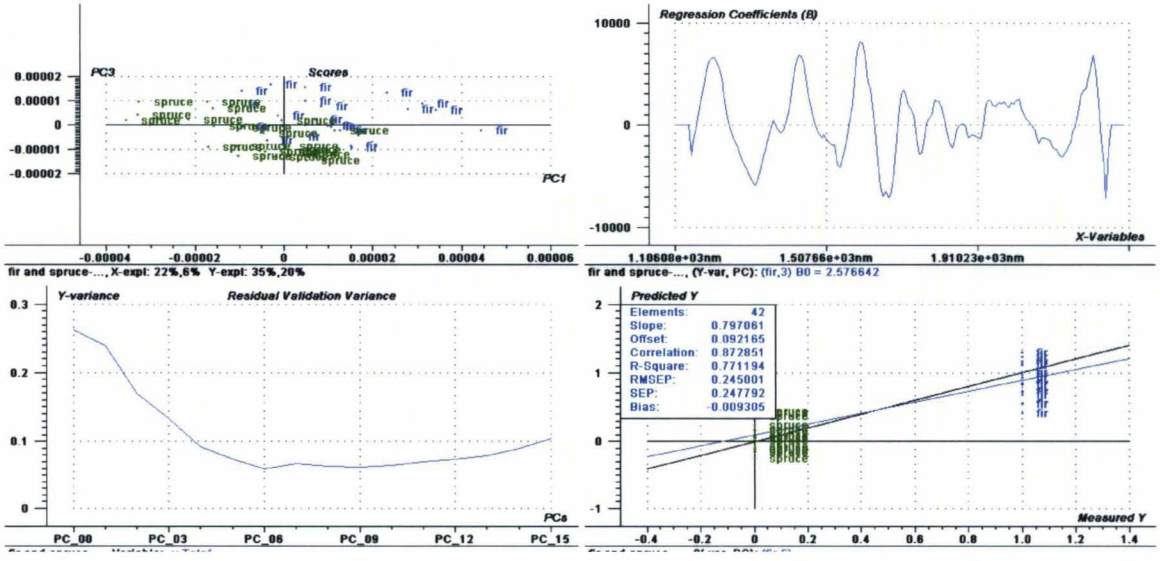


Fig. 3.26. PLS-DA regression validated model of fir and pine predictions subjected to various drying conditions: the upper left plot is Scores plot; the upper right plot is the Regression Coefficients plot for PC 3; the bottom left plot is the Residual Validation Variance plot; and the bottom right plot is the Predicted vs. Measured plot for fir with 8 PCs.



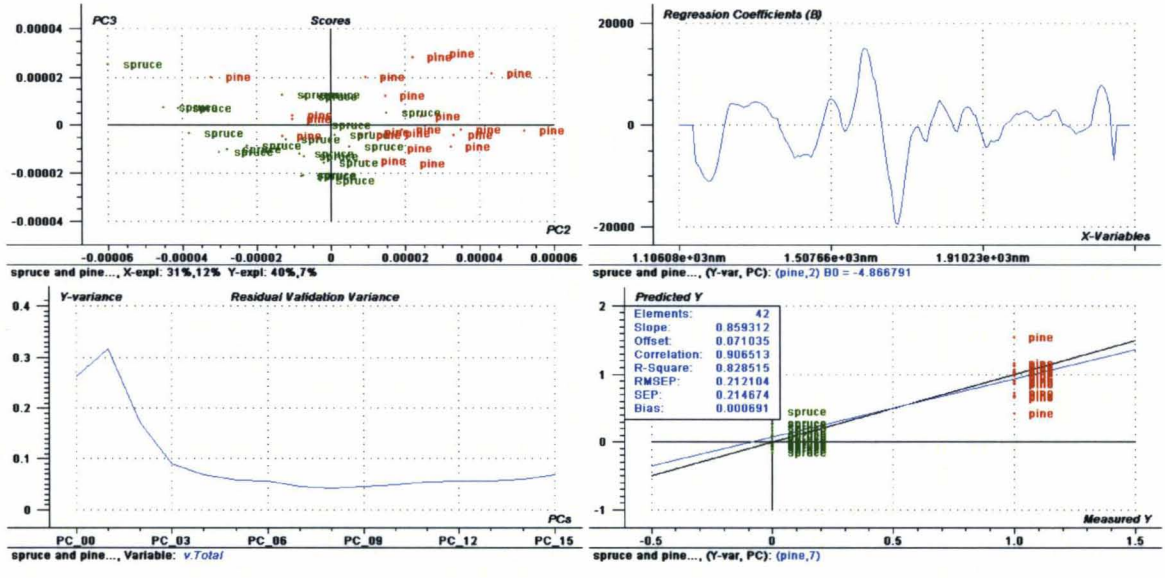
The regression coefficient plot for pine was the same but in opposite direction; and Predicted vs. Measured plot was the same for pine (not shown).

Fig. 3.27. PLS-DA regression validated model of fir and spruce predictions subjected to various drying conditions: the upper left plot is Scores plot; the upper right plot is the Regression Coefficients plot for PC 3; the bottom left plot is the Residual Validation Variance plot; and the bottom right plot is the Predicted vs. Measured plot for fir with 6 PCs.



The regression coefficient plot for spruce was the same but in opposite direction; and Predicted vs. Measured plot was the same for spruce (not shown).

Fig. 3.28. PLS-DA regression validated model of pine and spruce predictions subjected to various drying conditions: the upper left plot is Scores plot; the upper right plot is the Regression Coefficients plot for PC 2; the bottom left plot is the Residual Validation Variance plot; and the bottom right plot is the Predicted vs. Measured plot for pine with 7 PCs.



The regression coefficient plot for spruce was the same but in opposite direction; and Predicted vs. Measured plot was the same for spruce (not shown).

Fig. 3.29. PLS-DA regression prediction model of all species subjected to various drying conditions: the upper left plot is Predicted vs. Reference plot of fir with 11 PCs; the upper right plot is Predicted vs. Reference plot of pine with 11 PCs; and the bottom plot is the Predicted vs. Reference plot of spruce with 11 PCs.

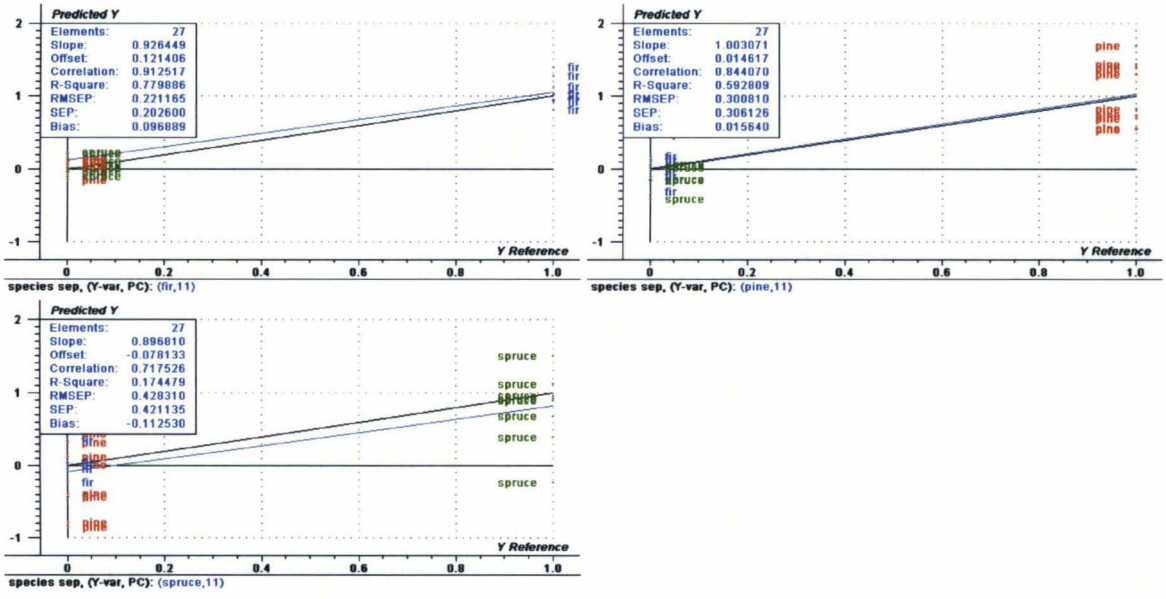


Fig. 3.30. PLS-DA regression prediction model of pine and fir subjected to various drying conditions: the upper plot is Predicted vs. Reference plot of pine with 8 PCs; and the bottom plot is the Predicted vs. Reference plot of fir with 8 PCs.

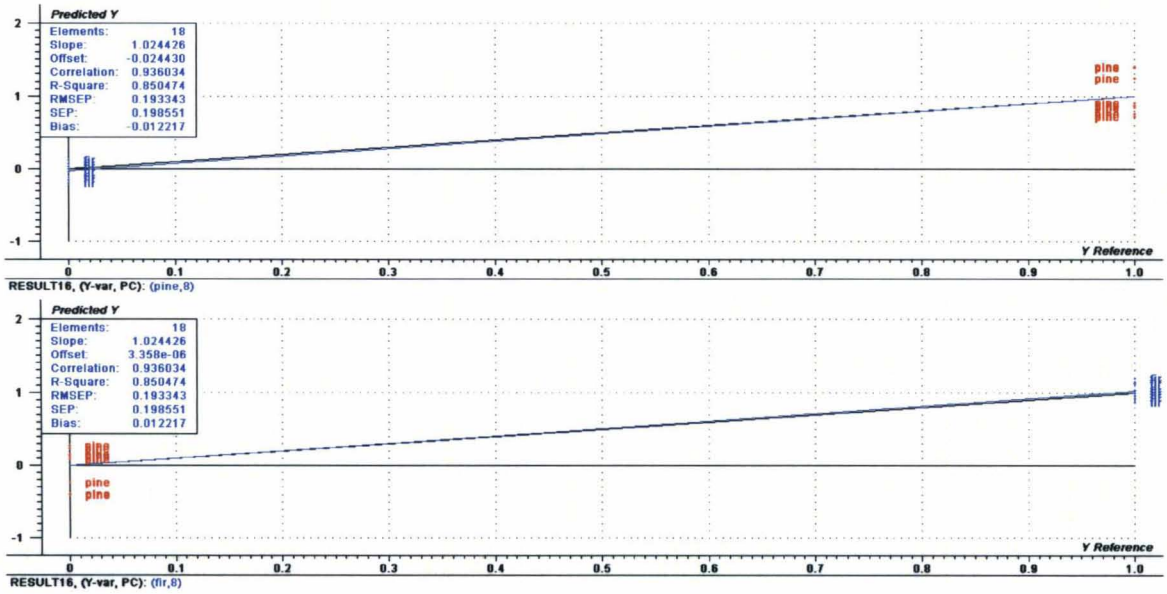


Fig. 3.31. PLS-DA regression prediction model of spruce and fir subjected to various drying conditions: the upper plot is Predicted vs. Reference plot of fir with 6 PCs; and the bottom plot is the Predicted vs. Reference plot of spruce with 6 PCs.

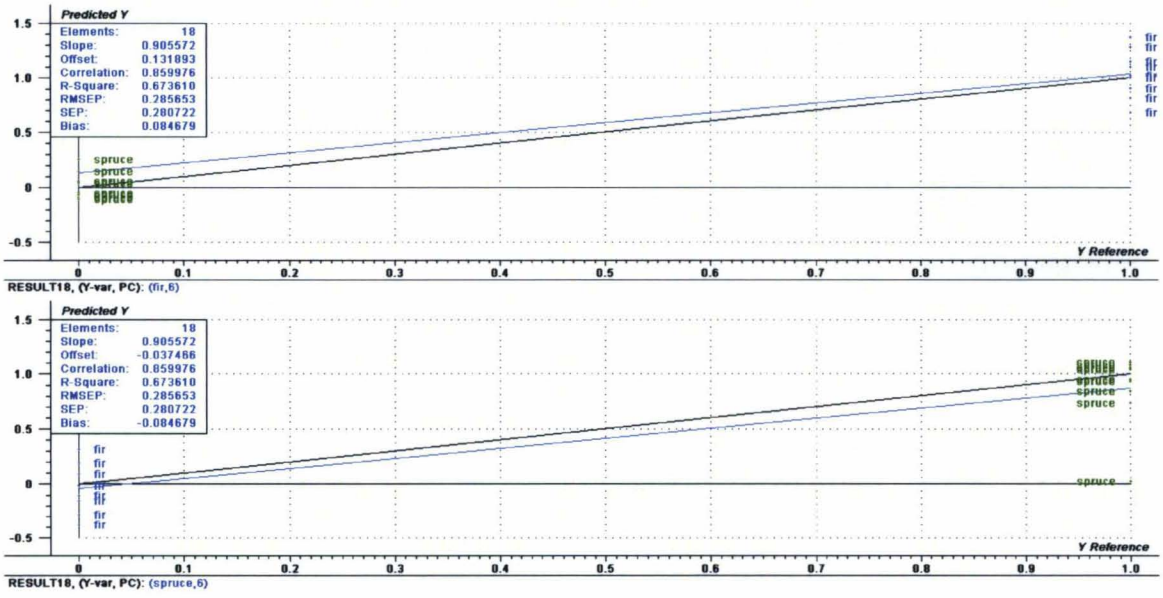


Table 3.5. Prediction results for the validated calibration PLS-DA regression models that predicted species subjected to various drying conditions.

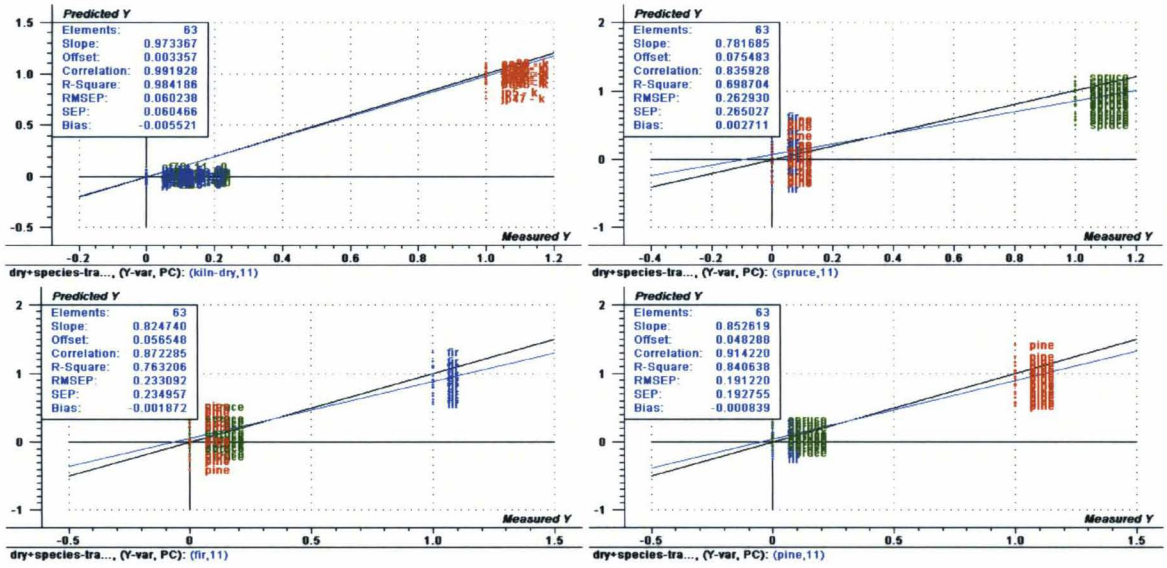
Species included in the model	**# of correct species type predictions (/9)**			**# PCs**	**R^2 (%)**			**RMSEP (/1)**		
	Pine	Spruce	Fir	Species	Pine	Spruce	Fir	Pine	Spruce	Fir
Spruce, fir, and pine	9	7*	9	11	59	17	78	0.3	0.43	0.22
Pine and fir	9	N/A	9	8	85	N/A	85	0.19	N/A	0.19
Fir and spruce	N/A	8**	9	6	N/A	67	67	N/A	0.28	0.28

* 1 spruce was assigned as fir and 1 spruce was not assigned to any of the species. **1 spruce was not assigned to any of the species.

A validated model that simultaneously predicted species and drying condition was developed (**Fig. 3.32**) and used to predict these variables. The calibration model with 11 PCs had R^2 and RMSEP of 84% and 19%, respectively for pine; 70% and 26% for spruce; 77% and 23% for fir; and 98% and 6% for drying conditions. No significant outliers were detected. Prediction results for drying conditions and species are seen in

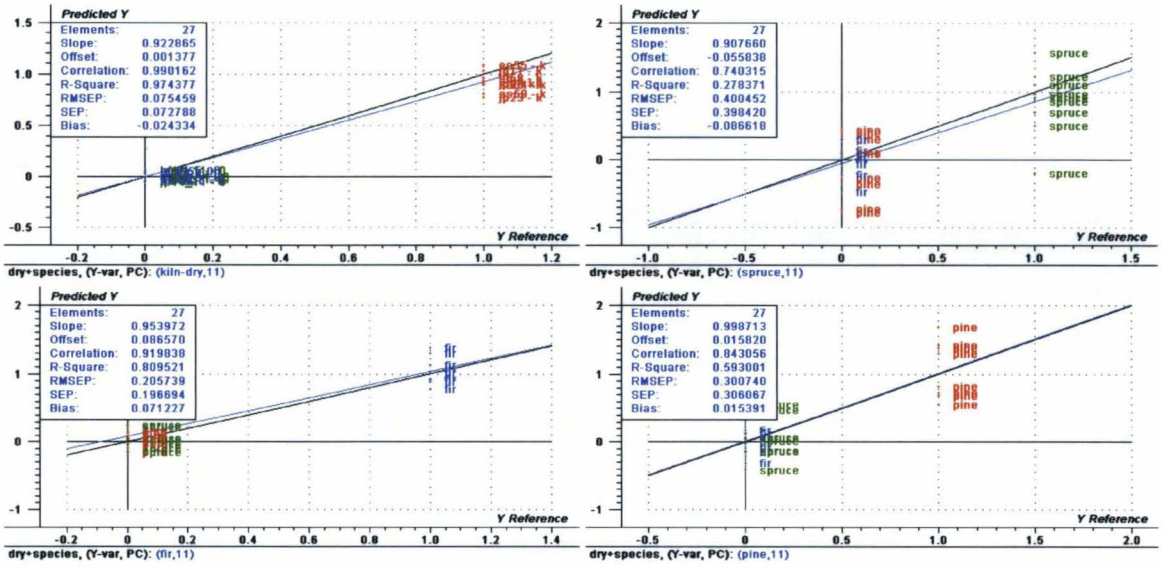
Fig. 3.33 and **Table 3.6**. Similar results were obtained when only two out of three species were used (**Fig. 3.34- 3.46** and **Table 3.6**). Despite the relatively low R^2 values and high RMSEP for species classification of models created on the training sets, the maximum error for species differentiation exhibited by the test sets was 7% (2 out of 27 samples were misclassified). In conclusion, both drying condition and species could be predicted simultaneously.

Fig. 3.32. PLS-DA regression validated model of both all species and drying conditions predictions: the upper left plot is Predicted vs. Measured plot for drying conditions predictions with 11 PCs; the upper right plot is the Predicted vs. Measured plot for spruce with 11 PCs; the bottom left plot is the Predicted vs. Measured plot for fir with 11 PCs; and the bottom right plot is the Predicted vs. Measured plot for pine with 11 PCs.



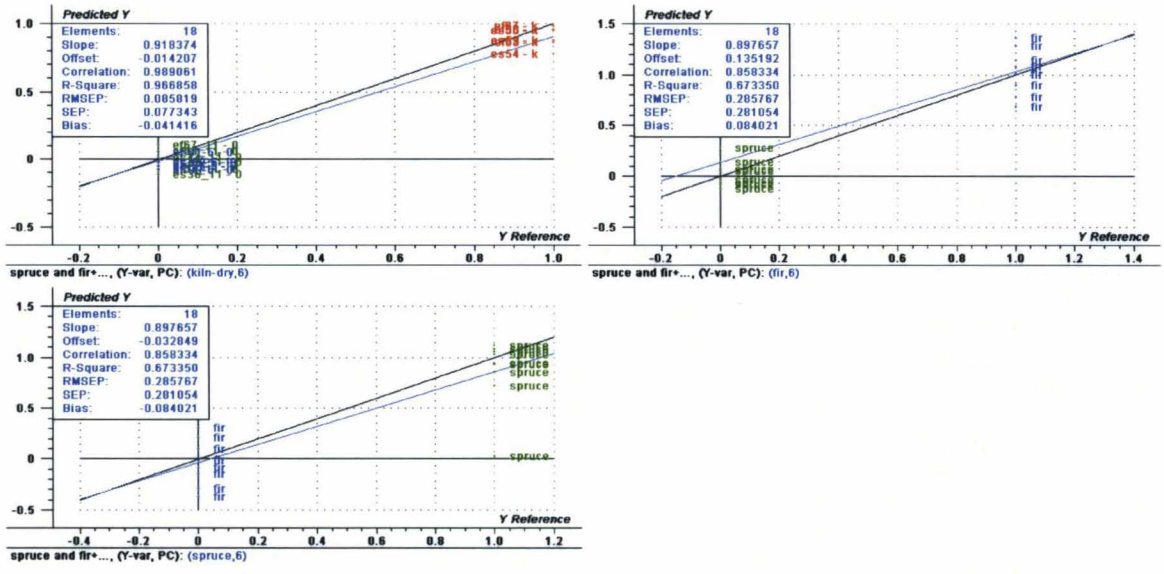
Blue-air-dried wood samples; green- 10% MC wood samples; and red-oven-dried wood samples. The Predicted vs. Measured plots were similar for air-dry and wet samples (not shown).

Fig. 3.33. PLS-DA regression prediction model of both all species and drying conditions: the upper left plot is Predicted vs. Reference plot for oven-dry prediction with 11 PCs; the upper right plot is the Predicted vs. Reference plot for spruce with 11 PCs; the bottom left plot is the Predicted vs. Reference plot for fir with 11 PCs; and the bottom right plot is the Predicted vs. Reference plot for pine with 11 PCs.



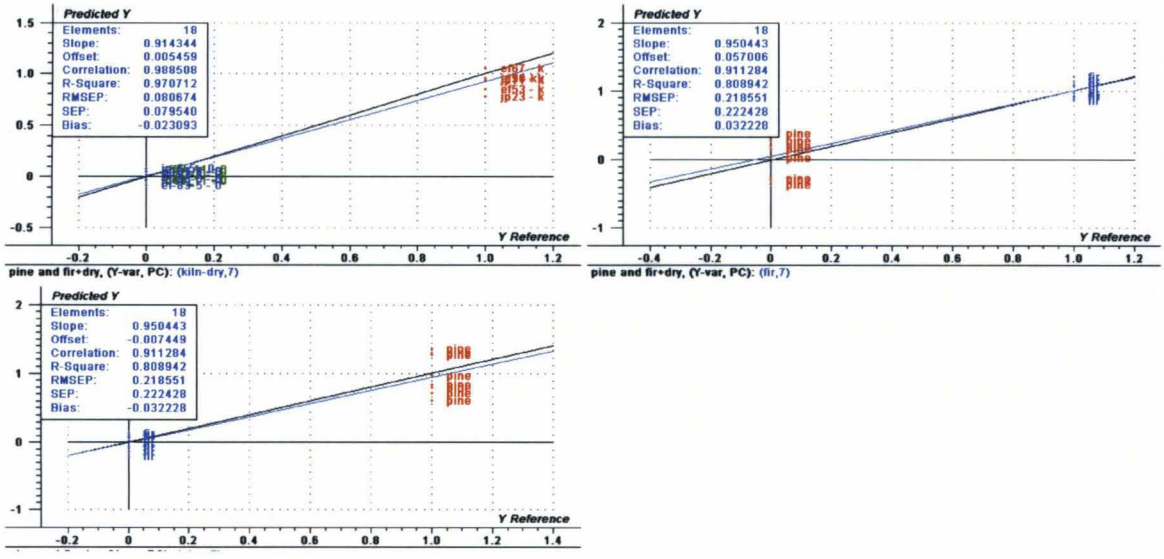
Blue- air-dried wood samples; green-10% MC wood samples; and red- oven-dried wood samples. The Predicted vs. Measured plots were similar for air-dry and wet samples (not shown).

Fig. 3.34. PLS-DA regression prediction model of both spruce and fir and drying conditions: the upper left plot is Predicted vs. Reference plot for drying condition prediction with 6 PCs; the upper right plot is the Predicted vs. Reference plot for fir with 6 PCs; and the bottom left plot is the Predicted vs. Reference plot for spruce with 6 PCs.



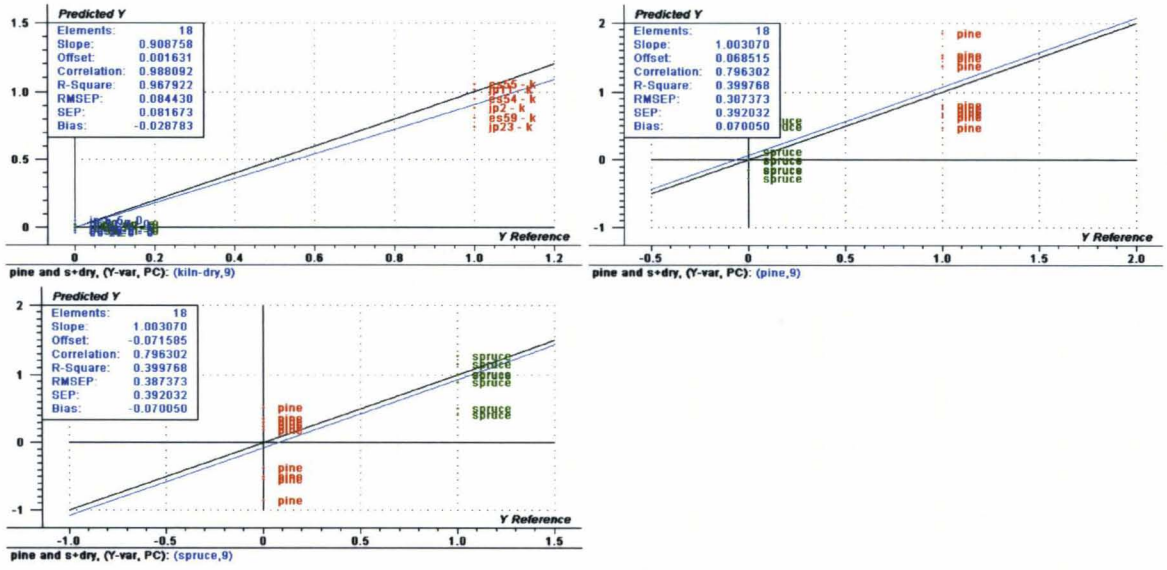
Blue- air-dried wood samples; green-10% MC wood samples; and red- oven-dried wood samples. The Predicted vs. Measured plots were similar for air-dry and wet samples (not shown).

Fig. 3.35. PLS-DA regression prediction model of both pine and fir and drying conditions: the upper left plot is Predicted vs. Reference plot for drying condition prediction with 7 PCs; the upper right plot is the Predicted vs. Reference plot for fir with 7 PCs; and the bottom left plot is the Predicted vs. Reference plot for pine with 7 PCs.



Blue- air-dried wood samples; green-10% MC wood samples; and red- oven-dried wood samples. The Predicted vs. Measured plots were similar for air-dry and wet samples (not shown).

Fig. 3.36. PLS-DA regression prediction model of both pine and spruce and drying conditions: the upper left plot is Predicted vs. Reference plot for drying condition prediction with 9 PCs; the upper right plot is the Predicted vs. Reference plot for pine with 9 PCs; and the bottom left plot is the Predicted vs. Reference plot for spruce with 9 PCs.



Blue- air-dried wood samples; green-10% MC wood samples; and red- oven-dried wood samples. The Predicted vs. Measured plots were similar for air-dry and wet samples (not shown).

Table 3.6. Prediction results for the validated calibration PLS-DA regression models that simultaneously predicted both species and drying conditions.

Variables included in the model	# of correct predictions (/9)				# PCs	R^2 (%)				RMSEP (/1)			
	Pine	Spruce	Fir	Oven-dry		Oven-dry	Pine	Spruce	Fir	Oven-dry	Pine	Spruce	Fir
Spruce, fir, and pine; air-dry, oven-dry, and wetter wood samples	9	8*	9	9	11	97	59	28	81	0.07	0.3	0.4	0.2
Pine and fir; air-dry, oven-dry, and wetter wood samples	9	N/A	9	9	7	97	81	N/A	81	0.08	0.22	N/A	0.22
Fir and spruce; air-dry, oven-dry, and wetter wood samples	N/A	8**	9	9	6	97	N/A	67	67	0.08	N/A	0.28	0.28
Pine and spruce; air-dry, oven-dry, and wetter wood samples	8***	8***	N/A	9	9	97	40	40	N/A	0.08	0.39	0.39	N/A

- 1 spruce was assigned as both fir and pine. **1 spruce was not assigned to any of the species. ***1 pine was assigned as spruce and 1 spruce was assigned as pine.
- Similar results to that of oven-dry were obtained for air-dry and wet wood samples (not shown).

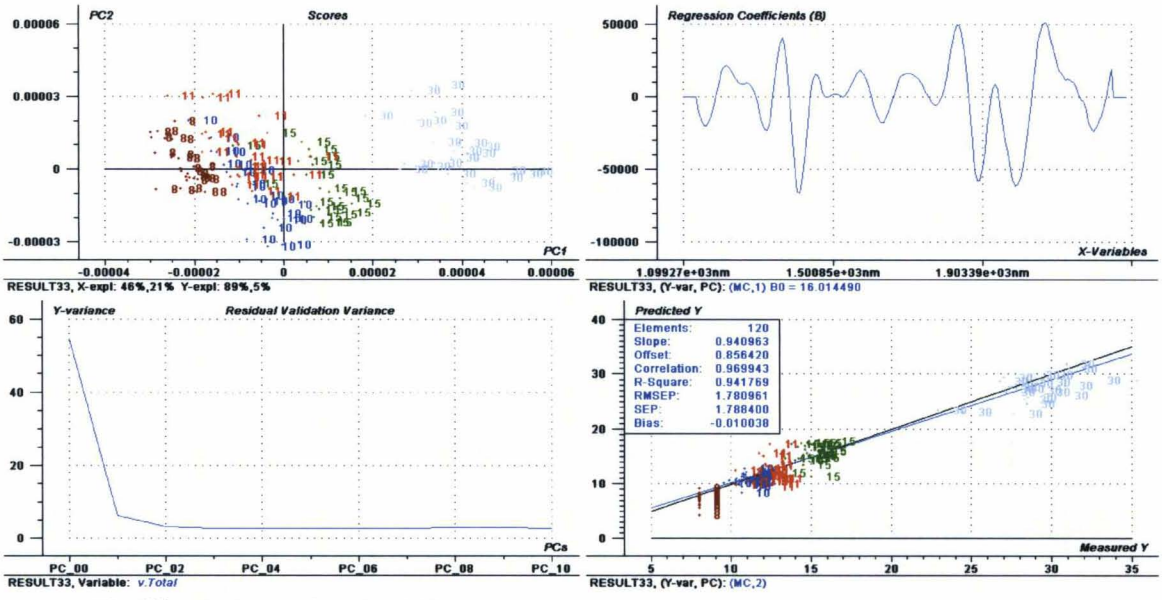
3.4.2. Moisture Content Analysis

Since the qualitative analysis worked well, a quantitative analysis was investigated. The same wood samples conditioned to different MC values (air-dry to FSP) were measured by NIR at specific equilibrium MCs. The treated absorbance spectra were then used for PLS analysis. **Fig. 3.37** shows the PLS-1 regression calibration validated model overview for MC separation of all species together. No outliers were detected and thus all samples were kept in the calibration model. Scores plot showed clustering of wood samples of the same or similar MCs with a clear distinction between samples of high moisture and low MC (**Fig. 3.37**). The calibration model with 2 PCs was a good model having R^2 and RMSEP of 94% and 5%, respectively (**Fig. 3.37**). In this model, only 46% variation in spectra explained 89% of variation in MC for the first PC (**Fig. 3.37**). The significant NIR wavelengths as seen in the regression coefficient plot were 1460, 1897, and 1992 nm (negative peaks) and 1365, 1470, 1835, and 2068 nm (positive peaks). **Fig. 3.38** shows the raw spectra of wood samples at MC of 8% and 30%. The absorption peak at 1460 and 1940 nm increased with increase in MC. These peaks are as a result of water vibrations. **Fig. 3.37**, on the other hand, mirrors the second derivative spectra with the absorption peaks pointing down rather than up. The second derivative spectrum is able to separate overlapping absorption bands because it is related to curvature and has the same sign as the curvature of a spectrum (4). For example, **Fig. 3.38** shows a single absorption band at 1460 and **Fig. 3.37** shows its second derivative. The left and right portions of the band are curved upward, so the second derivative is positive here (at 1365 and 1470 nm). The middle portion is curved downward, giving a negative second derivative with a

minimum at the peak value of the spectrum where the downward curvature is greatest (at 1460 nm). When these variables were used to construct a calibration model, the prediction error increased possibly because other variables played a role in MC prediction (plots not shown).

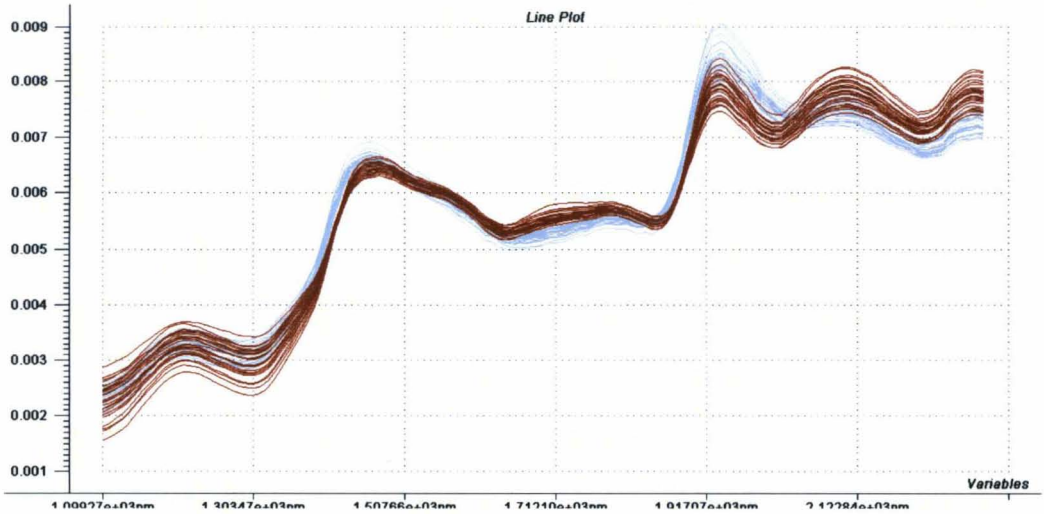
The calibration model was used to predict the MC values of the test samples. **Fig. 3.39** shows a plot of Predicted MC value using the calibration model vs. the laboratory measured MC, and a plot of Predicted MC with deviation vs. the test samples. The R^2 and RMSEP were 91% and 6%, respectively for a model with 2 PCs. Thus, the MC could be well predicted independently of the wood species analyzed.

Fig. 3.37. PLS-1 regression validated model of MC predictions of all species: the upper left plot is the Scores plot; the upper right plot is the Regression Coefficients plot for PC 1; the bottom left plot is the Residual Validation Variance plot; and the bottom right plot is the Predicted vs. Measured plot with 2 PCs.



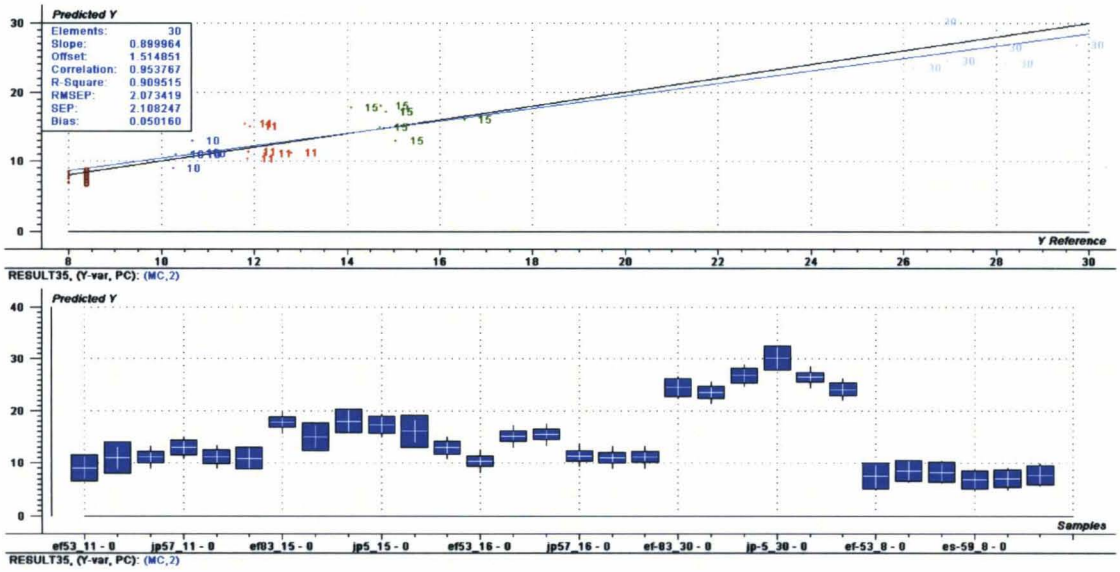
The different colors are the various MCs

Fig.3. 38. Raw spectra of wood samples conditioned to MC of 8% and 30%



Brown-wood samples at MC=8%; blue-wood samples at MC=30%

Fig. 3.39. PLS-1 regression prediction model of MC values of all species: the upper plot is Predicted vs. Reference plot with 2 PCs; and the bottom plot is the Predicted with deviation plot with 2 PCs.



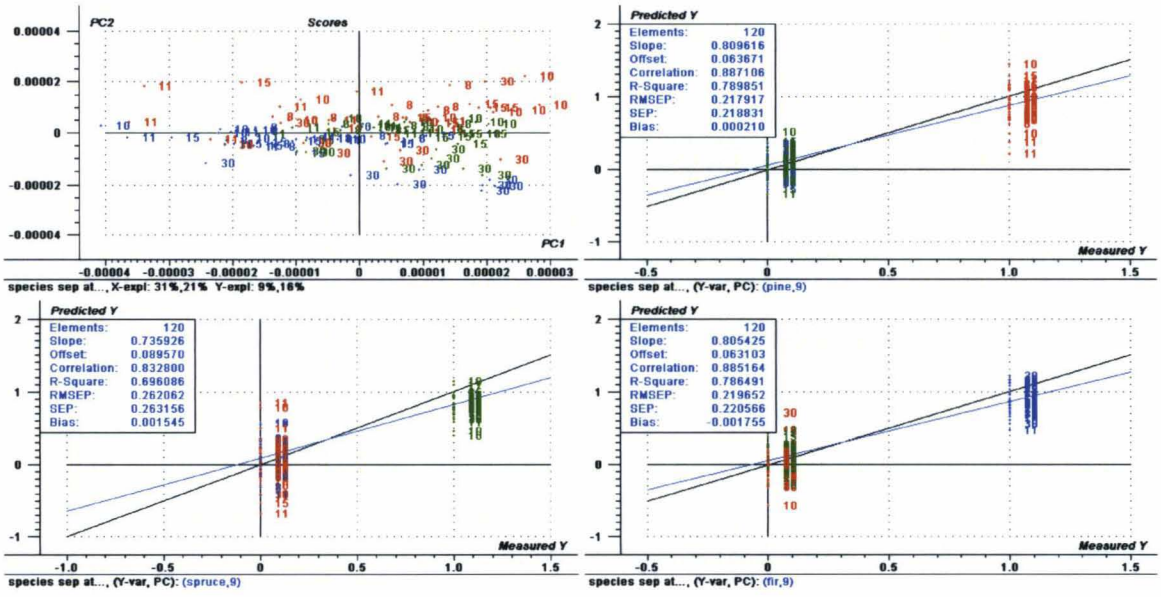
To test whether variation in MC could affect species separation, PLS-DA calibration validated models were constructed based on species separation at all MC values. When

all three species were included, the calibration model with 9 PCs (**Fig. 3.40**) had R^2 and RMSEP of 79% and 22%, respectively for pine; 70% and 26% for spruce; and 79% and 22% for fir. The scores plot showed no clear clustering but some separation of fir from pine by PC 2 was observed (**Fig. 3.40**). In this model, only 31% variation in spectra explained 9% of variation due to species differentiation for the first PC (**Fig. 3.40**). **Fig. 3.41** shows the significant variables needed for the three species separation (discussed in section 3.3) for PC 2 and separation of fir from pine was more obvious because of the strong inverse peaks of these species. Calibration models containing only two species were also constructed. They produced better models with more obvious clustering according to species and R^2 and RMSEP of 79% and 23%, respectively with 7 PCs for pine and spruce; 80% and 23% with 7 PCs for spruce and fir; and 94% and 12% with 9 PCs for pine and fir (**Fig. 3.42- 3.44**). No significant outliers were detected.

Fig. 3.45- 3.48 and **Table 3.7** include the prediction results for all the four models constructed. The best prediction models were obtained for fir and pine (R^2= 90%, RMSEP = 16% with 9 PCs) and fir and spruce (R^2 = 79%, RMSEP = 23% with 7 PCs) where all the species were determined correctly. In order to reduce the confounding effect of moisture, species differentiation was tested by creating a model without the variables required for MC separation. However, the prediction was essentially unchanged (plots not shown). The reason for this might be that change in MC caused changes along the full NIR spectrum of the samples, as shown in **Fig. 3.38**. Despite the relatively low R^2 values and high RMSEP of models created on the training sets, the maximum error for species

differentiation exhibited by the test sets was 10% (3 out of 30 samples were misclassified).

Fig. 3.40. PLS-DA regression validated model of all species predictions conditioned to various MC values: the upper left plot is the Scores plot; the upper right plot is the Predicted vs. Measured plot for pine with 9 PCs; the bottom left plot is the Predicted vs. Measured plot for spruce with 9 PCs; and the bottom right plot is the Predicted vs. Measured plot for fir with 9 PCs.



Red-pine; blue-fir; green-spruce

Fig. 3.41. Regression Coefficients of PC 2 from a PLS-DA regression validated model of species predictions conditioned to various MC values

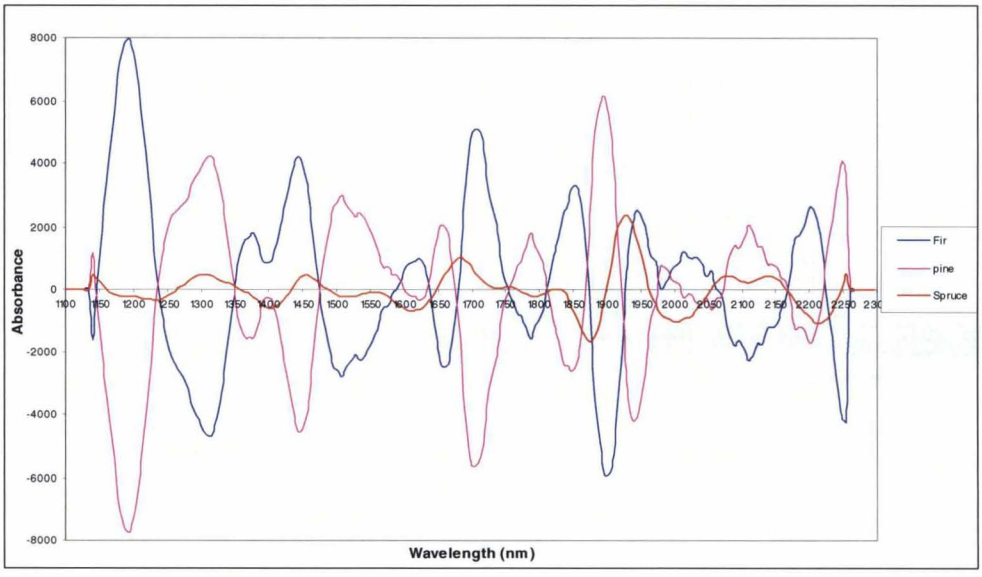
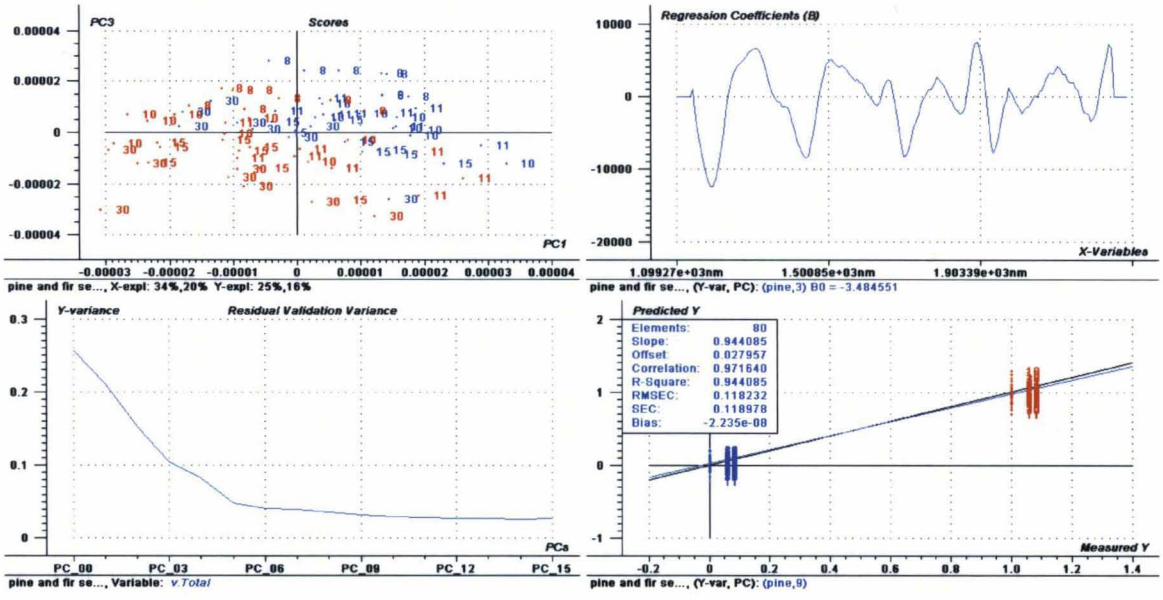
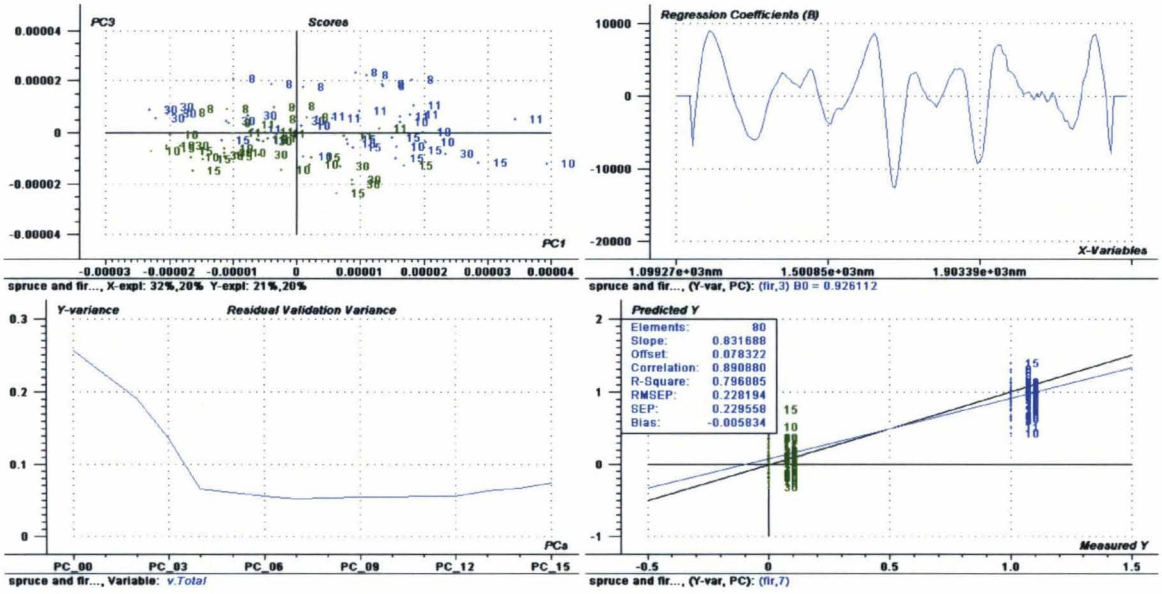


Fig. 3.42. PLS-DA regression validated model of pine and fir predictions conditioned to various MC values: the upper left plot is the Scores plot; the upper right plot is the Regression Coefficients plot for PC 3; the bottom left plot is the Residual Validation Variance plot; and the bottom right plot is the Predicted vs. Measured plot with 9 PCs.



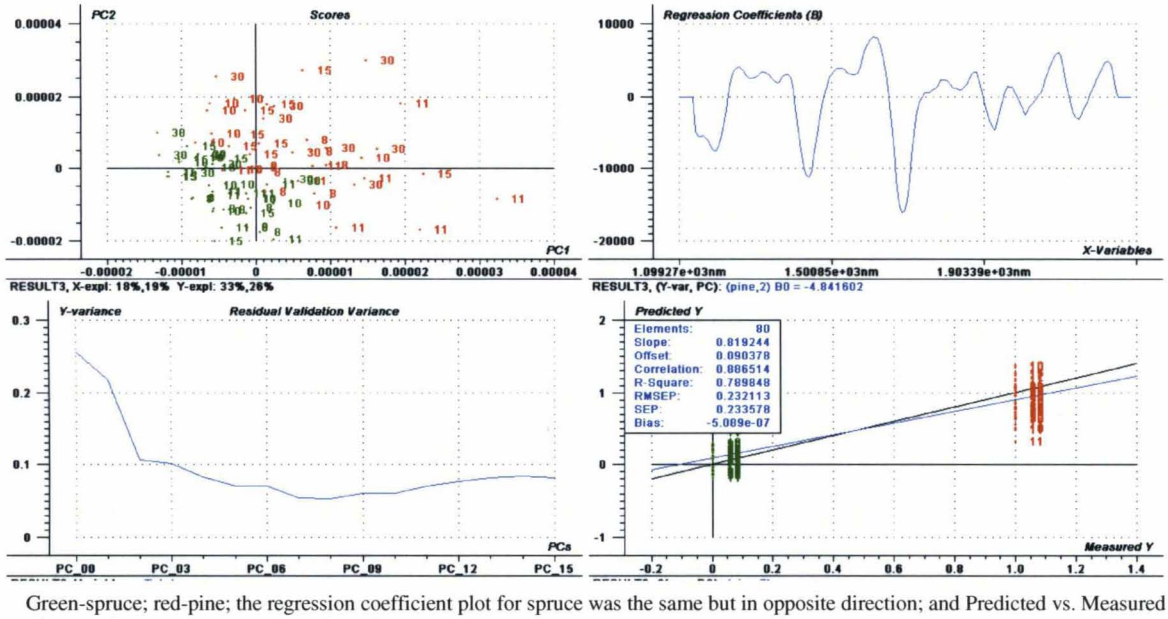
Red-pine; blue-fir; the regression coefficient plot for fir was the same but in opposite direction; and Predicted vs. Measured plot was the same for fir (not shown).

Fig.3. 43. PLS-DA regression validated model of spruce and fir predictions conditioned to various MC values: the upper left plot is the Scores plot; the upper right plot is the Regression Coefficients plot for PC 3; the bottom left plot is the Residual Validation Variance plot; and the bottom right plot is the Predicted vs. Measured plot with 7 PCs.



Green-spruce; blue-fir; the regression coefficient plot for spruce was the same but in opposite direction; and Predicted vs. Measured plot was the same for spruce (not shown).

Fig. 3.44. PLS-DA regression validated model of spruce and pine predictions conditioned to various MC values: the upper left plot is the Scores plot; the upper right plot is the Regression Coefficients plot for PC 2; the bottom left plot is the Residual Validation Variance plot; and the bottom right plot is the Predicted vs. Measured plot with 7 PCs.



Green-spruce; red-pine; the regression coefficient plot for spruce was the same but in opposite direction; and Predicted vs. Measured plot was the same for spruce (not shown).

Fig. 3.45. PLS-DA regression prediction model of all species conditioned to various MC values: the upper left plot is Predicted vs. Reference plot for pine with 9 PCs; the upper right plot is Predicted vs. Reference plot for fir with 9 PCs; and the bottom left plot is the Predicted vs. Reference plot for spruce with 9 PCs.

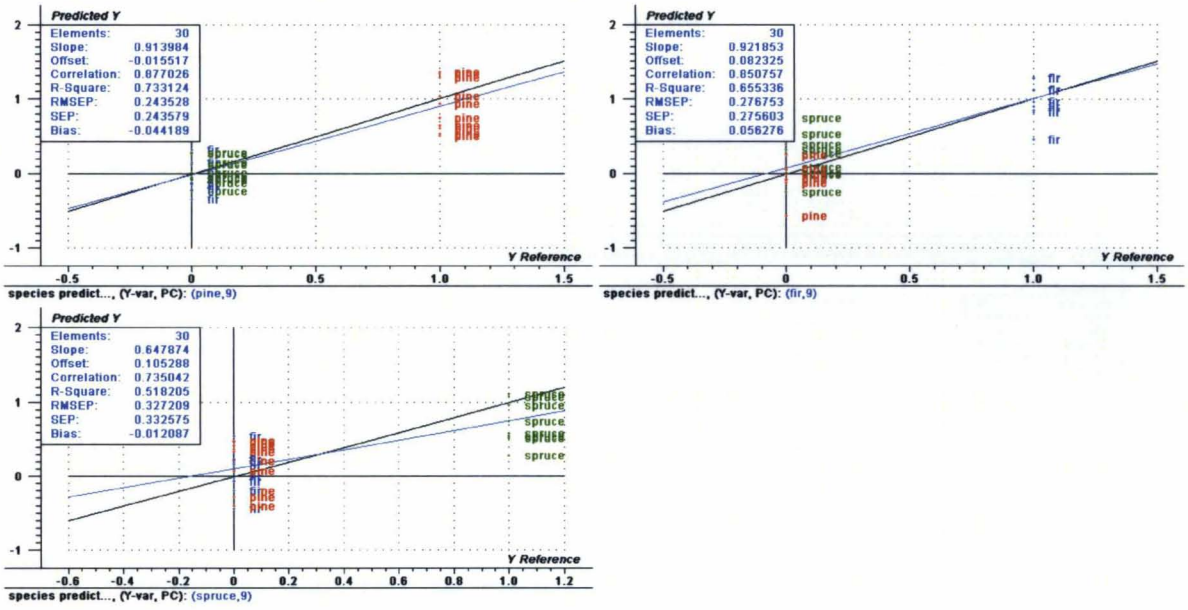


Fig. 3.46. PLS-DA regression prediction model of pine and fir conditioned to various MC values: the upper plot is Predicted vs. Reference plot for fir with 9 PCs; and the bottom plot is the Predicted vs. Reference plot for pine with 9 PCs.

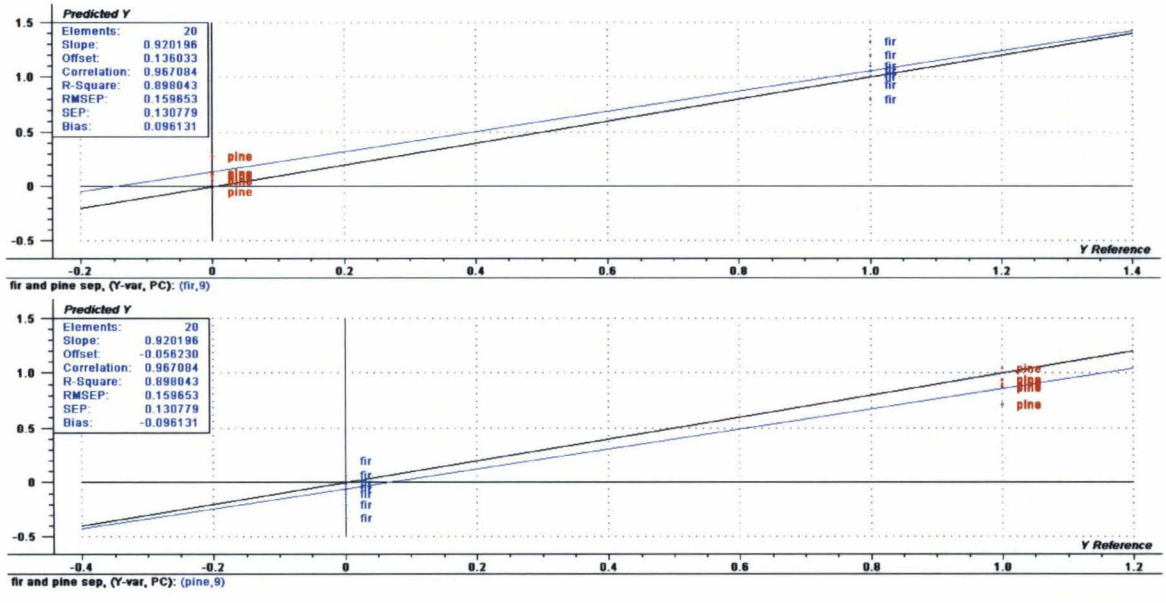


Fig. 3.47. PLS-DA regression prediction model of spruce and fir conditioned to various MC values: the upper plot is Predicted vs. Reference plot for spruce with 7 PCs; and the bottom plot is the Predicted vs. Reference plot for fir with 7 PCs.

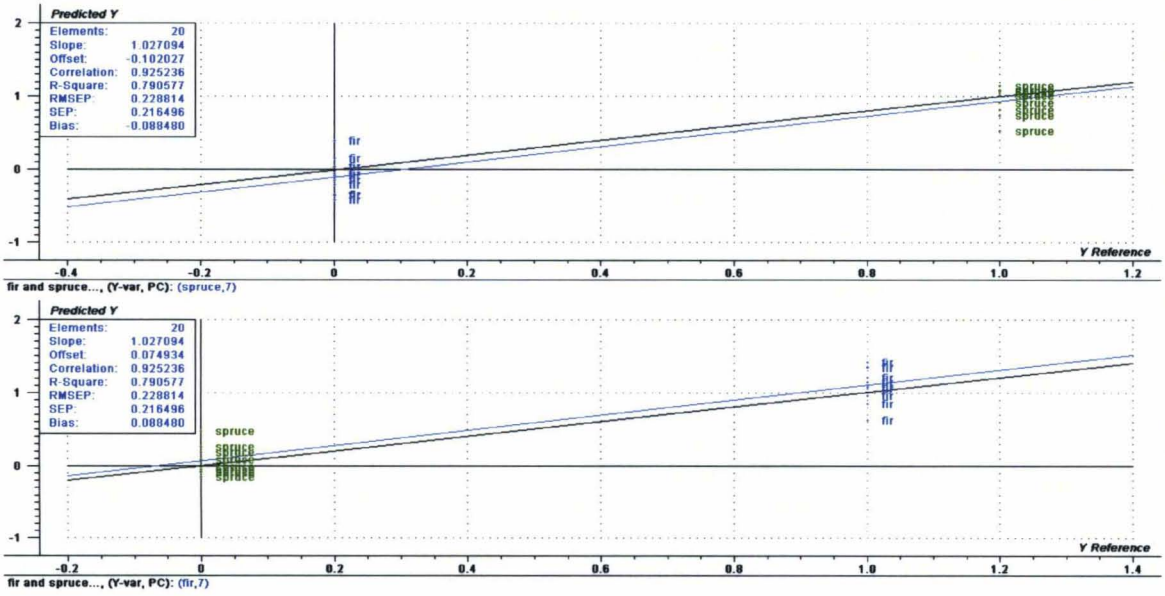


Fig. 3.48. PLS-DA regression prediction model of spruce and pine conditioned to various MC values: the upper plot is Predicted vs. Reference plot for spruce with 7 PCs; and the bottom plot is the Predicted vs. Reference plot for pine with 7 PCs.

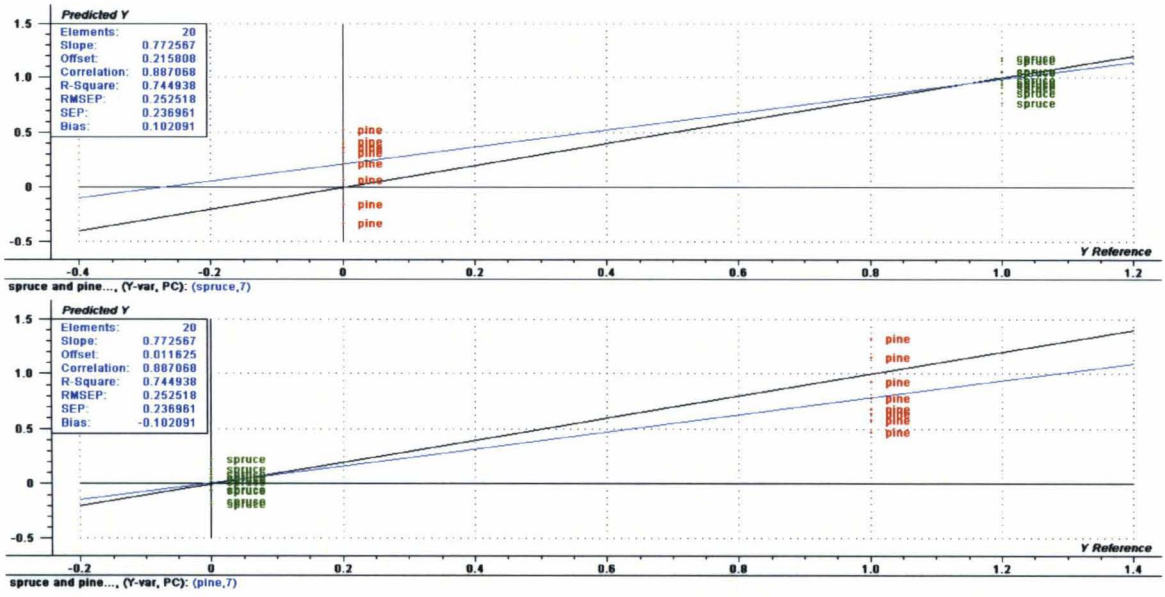


Table 3.7. Prediction results for all the validated calibration PLS-DA regression models of species conditioned to various MC values

Species included in the model	# of correct species predictions (/10)			# PCs	R^2 (%)			RMSEP (/1)		
	Pine	Spruce	Fir		Pine	Spruce	Fir	Pine	Spruce	Fir
Spruce, fir, and pine	10	8*	9*	9	73	52	65	0.24	0.33	0.28
Pine and fir	10	N/A	10	9	90	N/A	90	0.16	N/A	0.16
Fir and spruce	N/A	10	10	7	N/A	79	79	N/A	0.23	0.23
Pine and spruce	9**	10	N/A	7	74	74	N/A	0.25	0.25	N/A

* 2 spruces were assigned as fir and one fir was assigned as spruce. **1 pine was assigned as spruce.

A validated model that simultaneously predicted species and MC was developed (**Fig. 3.49**) and used to predict these variables. The calibration model containing all three species with 9 PCs had R^2 and RMSEP of 72% and 25%, respectively for pine; 55% and 32% for spruce; 67% and 27% for fir; and 93% and 6% for MC with 2 PCs. No significant outliers were detected. Prediction results for MC and species are seen in **Fig. 3.50** and **Table 3.8**. Similar prediction results were obtained when only two out of three species were used (**Fig. 3.51- 3.53** and **Table 3.8**) (calibration models not shown). Despite the relatively low R^2 values and high RMSEP for species classification of models created on the training sets, the maximum error for species differentiation exhibited by the test sets was 7% (2 out of 30 samples were misclassified). In conclusion, both MC values and species could be predicted simultaneously.

Fig. 3.49. PLS-DA regression validated model of both all species and MC predictions: the upper left plot is Predicted vs. Measured plot for pine with 9 PCs; the upper right plot is the Predicted vs. Measured plot for spruce with 9 PCs; the bottom left plot is the Predicted vs. Measured plot for fir with 9 PCs; and the bottom right plot is the Predicted vs. Measured plot for MC% prediction with 2 PCs.

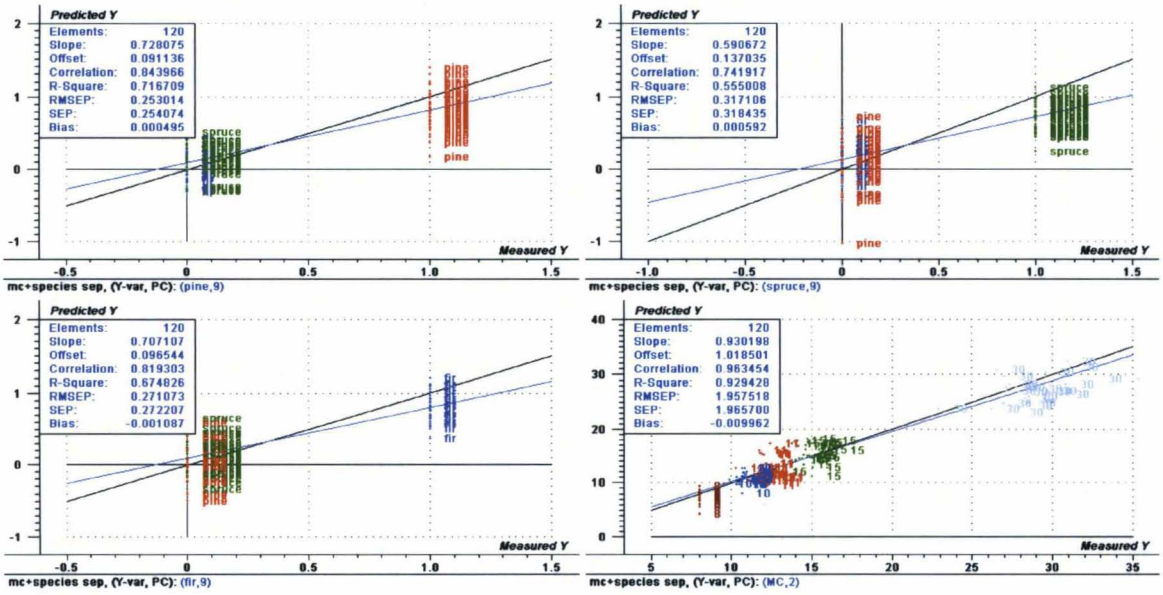


Fig. 3.50. PLS-DA regression prediction model of both all species and MC: the upper left plot is Predicted vs. Reference plot for MC% prediction with 2 PCs; the upper right plot is the Predicted vs. Reference plot for fir with 9 PCs; the bottom left plot is the Predicted vs. Reference plot for pine with 9 PCs; and the bottom right plot is the Predicted vs. Reference plot for spruce with 9 PCs.

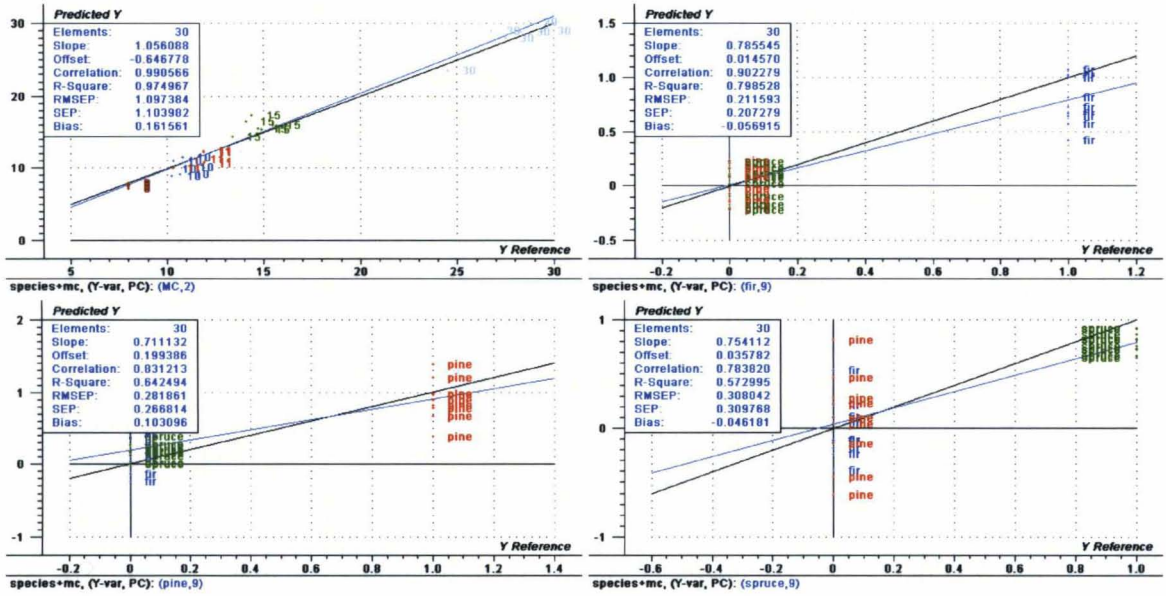


Fig. 3.51. PLS-DA regression prediction model of both spruce and fir and MC: the upper left plot is Predicted vs. Reference plot for MC% prediction with 2 PCs; the upper right plot is the Predicted vs. Reference plot for fir with 7 PCs; and the bottom left plot is the Predicted vs. Reference plot for spruce with 7 PCs.

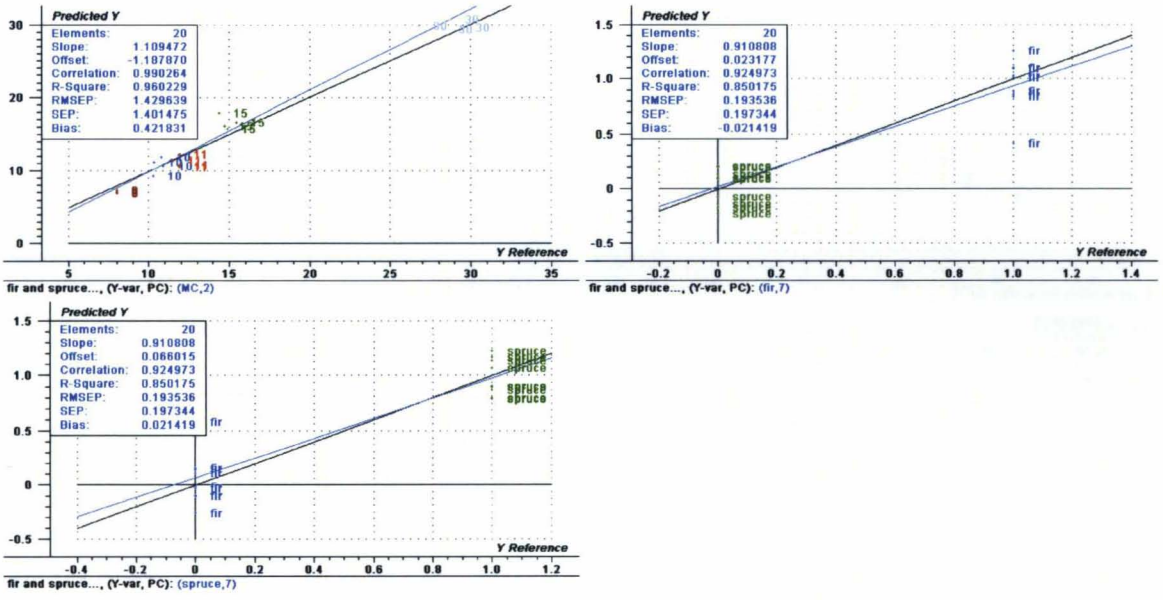


Fig. 3.52. PLS-DA regression prediction model of both pine and spruce and MC: the upper left plot is Predicted vs. Reference plot for MC% prediction with 2 PCs; the upper right plot is the Predicted vs. Reference plot for pine with 7 PCs; and the bottom left plot is the Predicted vs. Reference plot for spruce with 7 PCs.

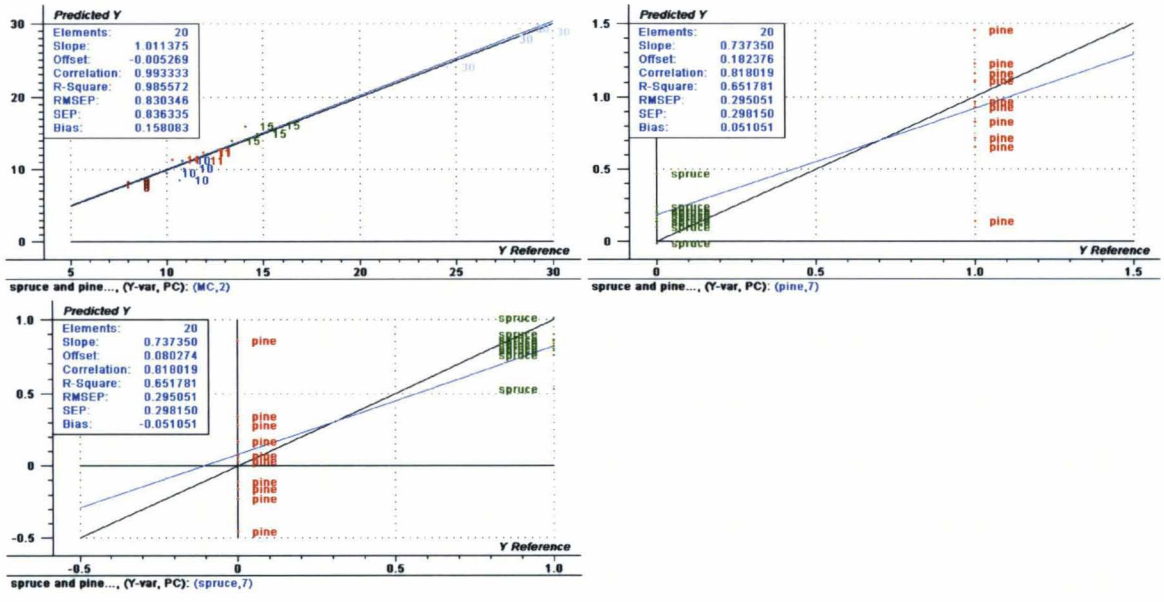


Fig. 3.53. PLS-DA regression prediction model of both fir and pine and MC: the upper left plot is Predicted vs. Reference plot for MC% prediction with 2 PCs; the upper right plot is the Predicted vs. Reference plot for fir with 9 PCs; and the bottom left plot is the Predicted vs. Reference plot for pine with 9 PCs.

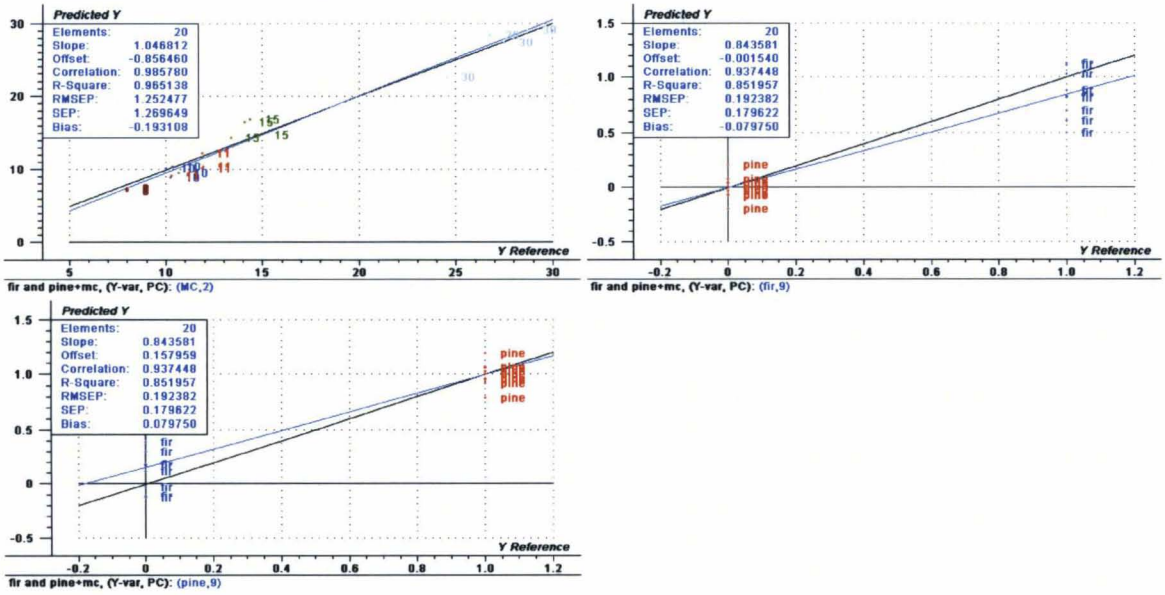


Table 3.8. Prediction results for the validated calibration PLS-DA regression models that simultaneously predicted both species and MC values.

Species included in the model	**# of correct species type predictions (/10)**			**# PCs**		**R^2 (%)**				**RMSEP (/35)**	**RMSEP (/1)**		
	Pine	Spruce	Fir	Species	MC%	MC%	Pine	Spruce	Fir	MC%	Pine	Spruce	Fir
Spruce, fir, and pine	9*	10	9*	9	2	97	64	57	80	1.1	0.28	0.31	0.21
Pine and fir	10	N/A	9**	9	2	96	85	N/A	85	1.25	0.19	N/A	0.19
Fir and spruce	N/A	10	9***	7	2	96	N/A	85	85	1.43	N/A	0.19	0.19
Pine and spruce	9****	10	N/A	7	2	98	65	65	N/A	0.83	0.29	0.29	N/A

* 1 pine and one fir were assigned as spruce. **1 fir was assigned as pine. ***1 fir was assigned as spruce. ****1 pine was assigned as spruce.

3.5. Wood Weathering Analysis

Weathering is the general term used to define the slow degradation of materials exposed to the weather. Wood weathering is surface degradation of wood that is initiated primarily by solar radiation, but other factors are also important. The wetting and drying of wood through precipitation, diurnal and seasonal changes in RH, temperature changes, oxygen, and human activities all contribute to the degradation of wood surfaces.

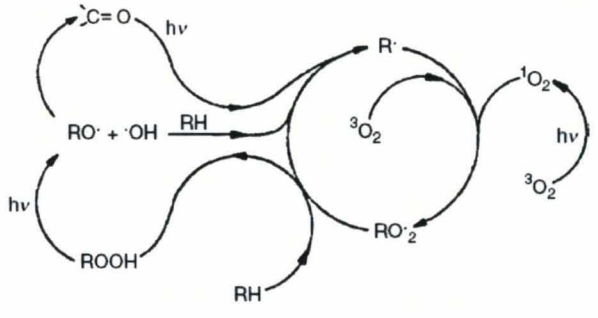
It is primarily the UV portion of the solar spectrum that initiates the process of weathering. It is a photooxidation or photochemical degradation of the wood surface. The UV radiation has sufficient energy to chemically degrade wood structural components (primarily lignin). The degradation starts immediately after the wood is exposed to sunlight. First the color changes, then the surface fibers loosen and erode, but the process is rather slow. It can take more than 100 years of weathering to decrease the thickness of a board by 5-6 mm. (5)

The UV and Vis solar radiation that reaches the earth's surface is limited to the range between 295–400 nm and 400-800 nm, respectively. In order for a photochemical reaction to occur, sufficient energy to disrupt a chemical bond (bond dissociation energy) must be absorbed by some chemical moiety in the system. Using equation [2], the energy for UV radiation at a wavelength of 295 nm is about 97 Kcal/mole and for 400 nm is about 72 Kcal/mole. Several of the chemical moieties have bond dissociation energies well above the energy of terrestrial UV radiation and therefore cannot be affected by

natural UV radiation. The bond dissociation energy must be below 97 kcal/mole for the chemical moiety to absorb radiation. The bond dissociation energies for many of the carbon-oxygen moieties commonly found in lignin fall within the UV radiation range. (5)

There have been many studies to investigate the mechanism of wood weathering, and it has been clearly shown that the absorption of a UV photon can result in the formation of a free radical, and that through the action of oxygen and water, a hydroperoxide is formed. Both the free radical and hydroperoxide can initiate a series of chain scission reactions to degrade the polymeric components of wood. Despite many studies spanning several decades, the mechanism is still not well defined and can only be represented in a general way (**Fig. 3.54**). (5)

Fig. 3.54. Mechanism of photodegradation of wood (3)



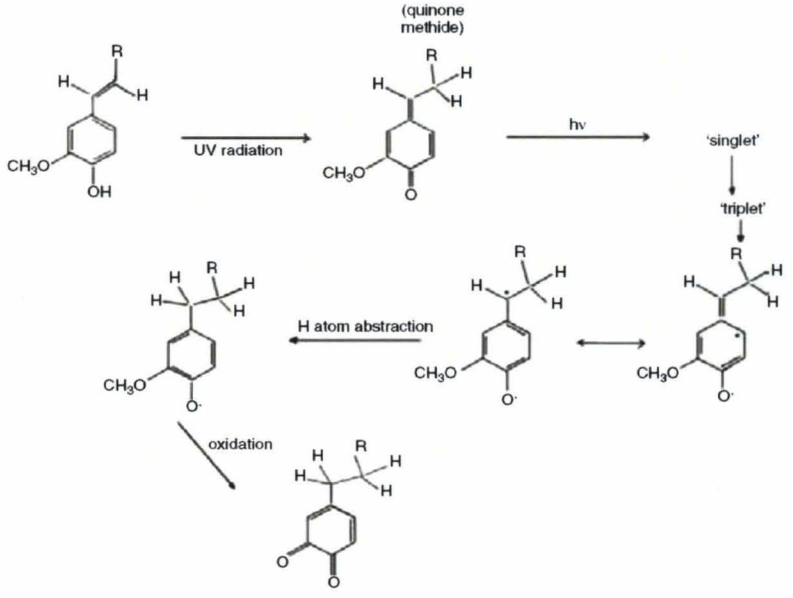
By comparing the energy available from the photons in the UV range of the spectrum, it is apparent that there is sufficient energy to break bonds in the chemicals that comprise wood. However, in order for a bond to break, energy must be absorbed by some component of the wood. The absorbed energy promotes the molecule to an excited state

that can be dissipated through a number of paths. The most benign would be a return to the ground state through dissipation of heat. Other alternatives would involve chemical reactions. One of the common chemical reaction paths following the absorption of a quantum of energy is chemical dissociation to form a free radical. Since wood does not normally have free radicals, their presence following UV irradiation signals the dissociation of a chemical bond of lignin moieties having α-carbonyl, biphenyl, or ring-conjugated double bond structures. From work on several lignin model compounds, Hon (88) demonstrated that the phenoxy radicals produced from phenolic hydroxyl groups by the action of light were the major intermediates formed in the photodegradation of lignin, and that these intermediates reacted with oxygen and demethylated to form an *o*-quinonoid structure. Free radicals generated in wood are also thought to react with molecular oxygen to form hydroperoxides. Additional reactions result in the formation of carboxyls, carbonyls, quinones, and loss of lignin. (88)

The degradation of wood by UV light results in the reduction in the methoxyl and lignin content of wood, and an increase in its acidity and carbonyl content. IR spectroscopic characterization of UV-exposed wood surfaces has shown an increase in carbonyl and carboxylic functional groups, and a decrease in aromatic functional groups due to degradation of lignin. However, there can also be a decrease in carbonyl absorption due to leaching of the surface by rain/water as the exposure progresses. (5) Anderson (89) used a xenon light source to approximate natural UV radiation. They measured the surface degradation over 2400 hours of UV light exposure with daily water spray of 4 hours, or light without the water spray. The surface degradation of softwoods

was evaluated using reflectance FT-IR. The carbonyl absorption increased in intensity during the early exposure then decreased (removed from the surface), and the aromatic absorption decreased. All softwoods had a distinctly cellulosic spectra following 2400 hours of accelerated weathering (light and water), indicating a loss of lignin. A mechanism was proposed (**Fig. 3.55**).

Fig. 3.55. Lignin photooxidation mechanism (89)



60 samples (20 of each species of SPF) were artificially weathered for 0-8 weeks and then measured by NIRS every two weeks. Their transformed absorbance spectra were compared by using PLS-1. The validated calibration model is shown in **Fig. 3.56** and the prediction results in **Fig. 3.57**. The scores plot showed clustering of wood samples according to weathering period by PC 1 (**Fig. 3.56**). The calibration model with 6 PCs was moderate with R^2 and RMSEP of 87% and 12%, respectively which could be explained by non-linear changes with weathering time (**Fig. 3.56**). In this model, only

28% variation in spectra explained 67% of variation due to weathering period of wood samples for the first PC (**Fig. 3.56**).

The significant variables needed for the separation due to weathering periods were at 1960, 2102, and 2199 nm (negative peaks); 1460, 1678, 2006, 2288nm (positive peaks) (**Fig. 3.56**). **Fig. 3.58** shows the raw spectra of non-weathered and weathered (2 and 8 weeks) wood samples. The absorption band at 1678 nm which was assigned to CH stretch first overtone due to the aromatic portion of lignin (2) decreased with weathering period as a result of lignin degradation. The absorption band at 1960 nm which first increased with weathering and then decreased was attributed to the C=O stretch second overtone (2). The carbonyl absorption increased in intensity during the early exposure then decreased because of removal of lignin from the surface by water. This explains why the wood bleached eventually. The absorption bands at 1460 and 1931 and 2006 nm which first increased and then decreased were assigned to the OH stretch first overtone and the OH stretch and OH deformation combination, respectively (2). These absorptions first increased possibly due to soluble extractives which were brought by the water to the wood surface (explaining why the wood was darker at the early exposure times) and then they decreased due to their sequential removal from the surface probably together with cellulose and/or hemicellulose degradation. Degradation of hemicellulose could also be supported by the fact that the absorption band at 1712 nm which was attributed to the CH stretch first overtone due to hemicellulose (2) decreased with weathering. The absorption band at 2288 nm which was assigned to the CH stretch and CH deformation of methyl group (2) decreased with weathering as a result of lignin degradation. Finally, the

absorption bands at 2102 and 2198 nm were assigned to the CH stretch and C=O stretch (2) and thus, no obvious trend was seen. The samples in the test set were predicted based on weathering period with R^2 and RMSEP of 85% and 12%, respectively for 6 PCs independently of the wood species analyzed (**Fig. 3.57**). No outliers were detected.

Fig. 3.56. PLS-1 regression validated model of weathering period predictions: the upper left plot is Scores plot; the upper right plot is the Regression Coefficients plot for PC 1; the bottom left plot is the Residual Validation Variance plot; and the bottom right plot is the Predicted vs. Measured plot with 6 PCs.

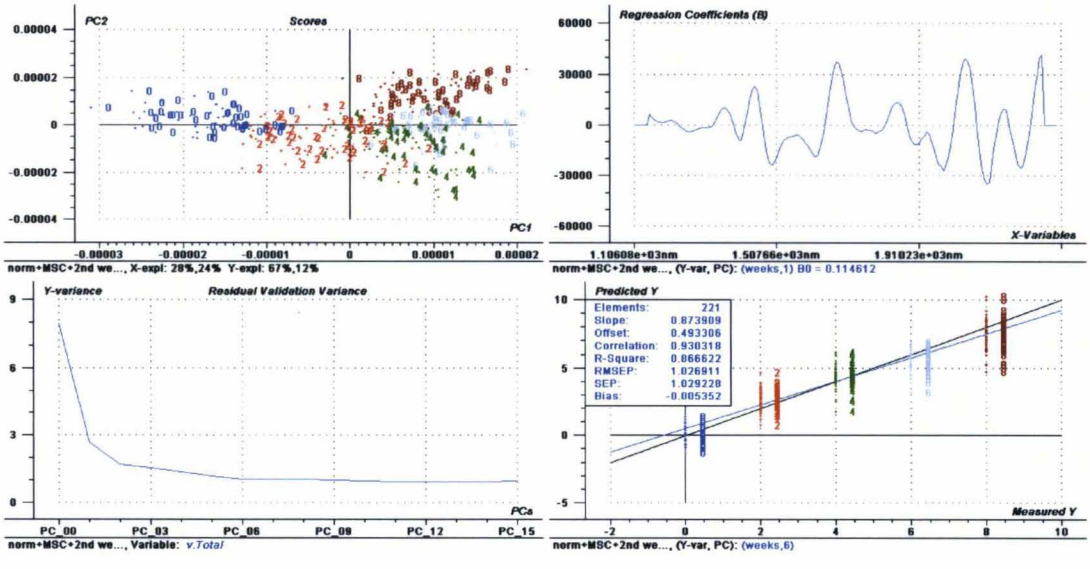


Fig. 3.57. PLS-1 regression prediction model of weathering periods of all species: the upper plot is Predicted vs. Reference plot with 6 PCs; the; and the bottom plot is the Predicted with deviation plot.

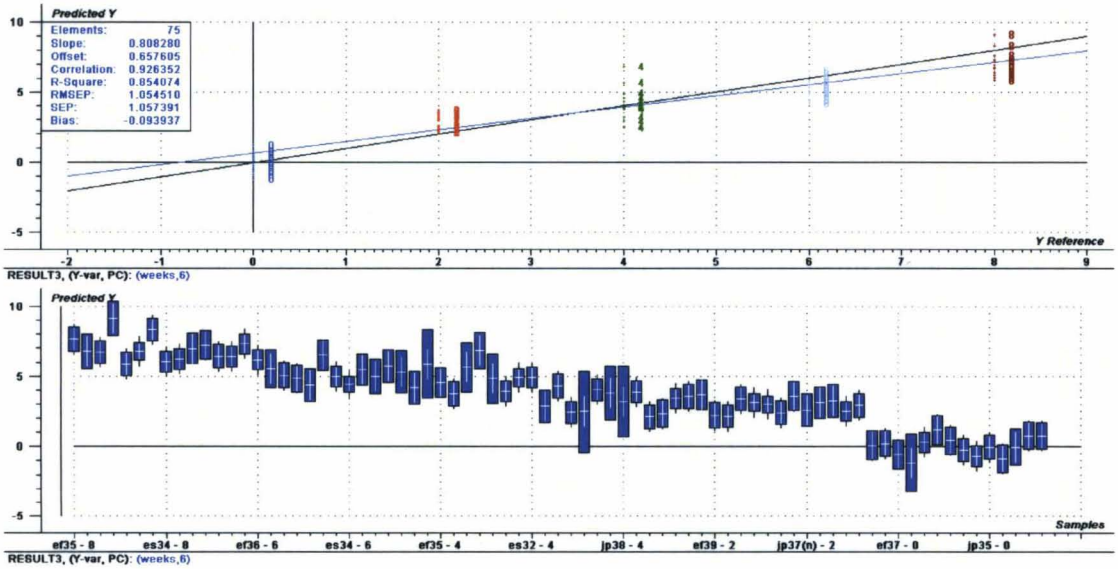
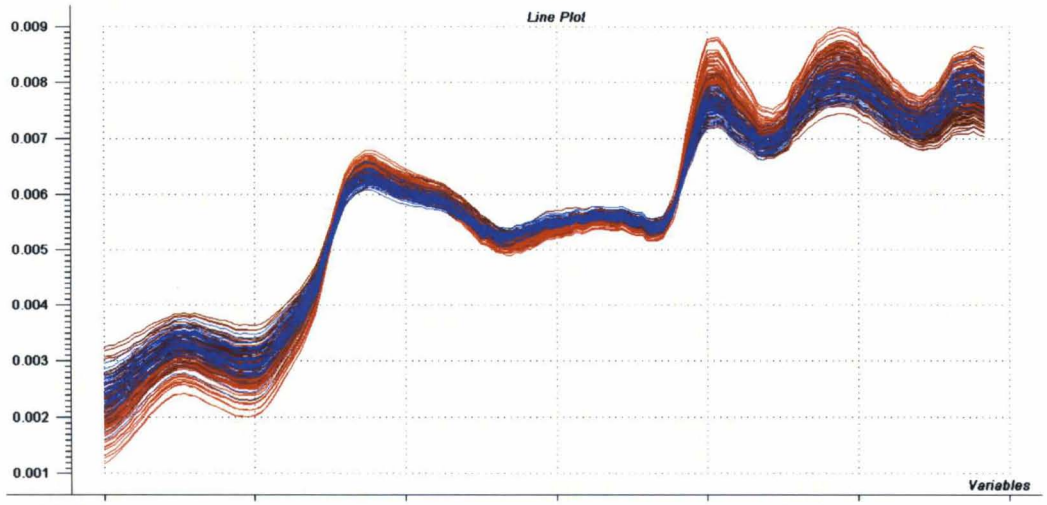


Fig.3.58. Raw spectra of non-weathered and weathered wood samples



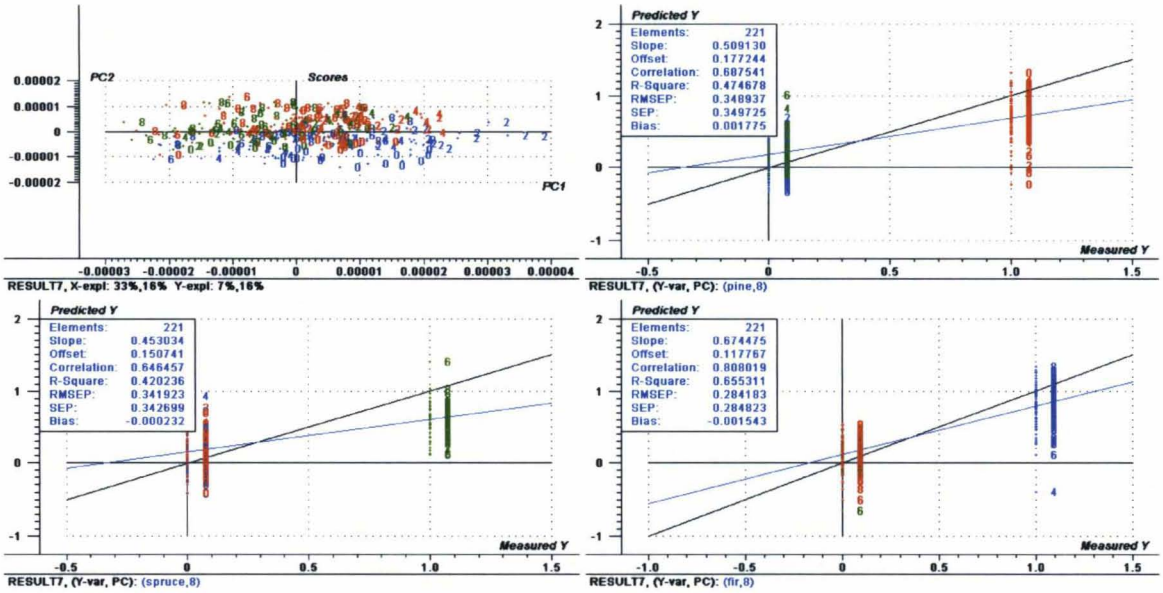
Red- wood samples weathered for 2 weeks; blue-non-weathered wood samples; and brown- wood samples weathered for 8 weeks.

It was then determined whether surface changes due to weathering could affect species separation. To investigate this, PLS-DA calibration validated models were constructed based on species separation at all weathering periods. When all three species were included, a relatively poor calibration model based on species classification with 8 PCs was obtained (**Fig. 3.59**) and R^2 and RMSEP were 47% and 35%, respectively for pine; 42% and 34% for spruce; and 65% and 28% for fir. The scores plot showed no clear clustering but some separation of fir from pine and spruce by PC 2 was observed (**Fig. 3.59**). In this model, only 33% variation in spectra explained 7% of variation due to species differentiation for the first PC (**Fig. 3.59**). **Fig. 3.60** shows the significant variables needed for the three species separation (discussed in section 3.3) for PC 2 and again separation of fir from pine and spruce was more obvious because of the strong inverse peaks of these species. Calibration models containing only two species were also constructed. They produced better models with more obvious clustering according to species and R^2 and RMSEP of 64% and 29%, respectively with 6 PCs for pine and fir; 65% and 29% with 6 PCs for spruce and fir; and 51% and 35% with 10 PCs for spruce and pine (**Fig. 3.61- 3.63**). No significant outliers were detected.

Fig. 3.64- 3.67 and **Table 3.9** include the prediction results for all the models constructed. The best prediction model was obtained for fir and spruce (R^2 = 66%, RMSEP = 29% with 6 PCs) where all the species except 2 out of 48 were determined correctly. In order to reduce the confounding effect of weathering, species differentiation was tested by creating a model without the variables required for separation due to weathering. However, the prediction was essentially unchanged (plots not shown). The

reason for this might be that variation due to weathering caused changes along the full NIR spectrum of the samples, as shown in **Fig. 3.58**. Despite the relatively low R^2 values and high RMSEP of models created on the training sets, the maximum error for species differentiation exhibited by the test sets was 8% (6 out of 75 samples were misclassified).

Fig. 3.59. PLS-DA regression validated model of all species predictions whose surface was weathered for various periods: the upper left plot is Scores plot; the upper right plot is the Predicted vs. Measured plot for pine with 8 PCs; the bottom left plot is the Predicted vs. Measured plot for spruce with 8 PCs; and the bottom right plot is the Predicted vs. Measured plot for fir with 8 PCs.



Red-pine; blue-fir; green-spruce

Fig. 3.60. Regression Coefficients of PC 2 from a PLS-DA regression validated model of species predictions whose surface was weathered for various periods.

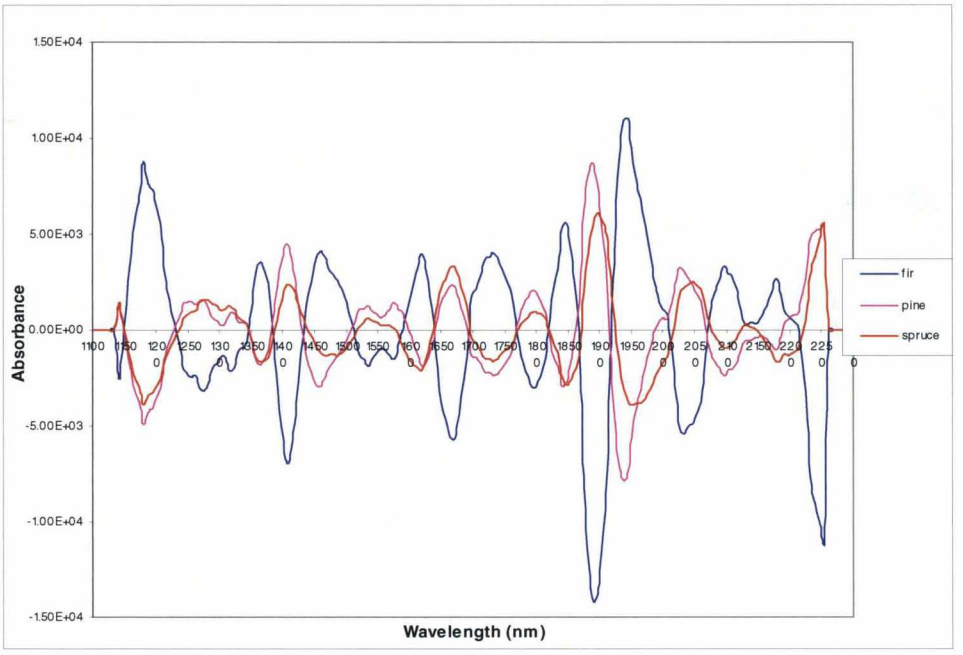
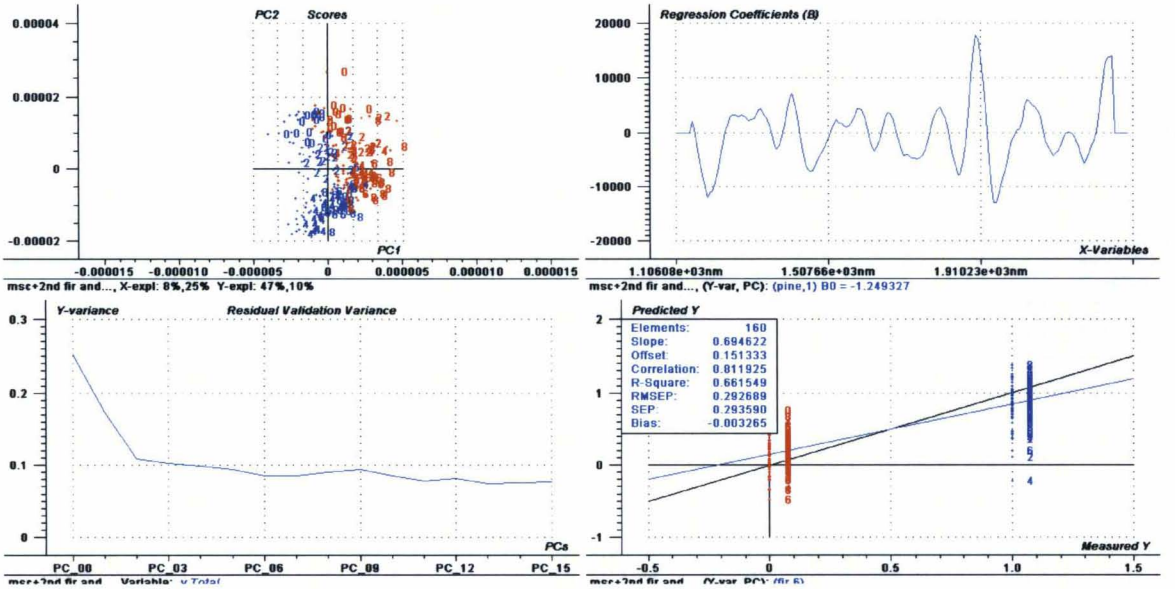
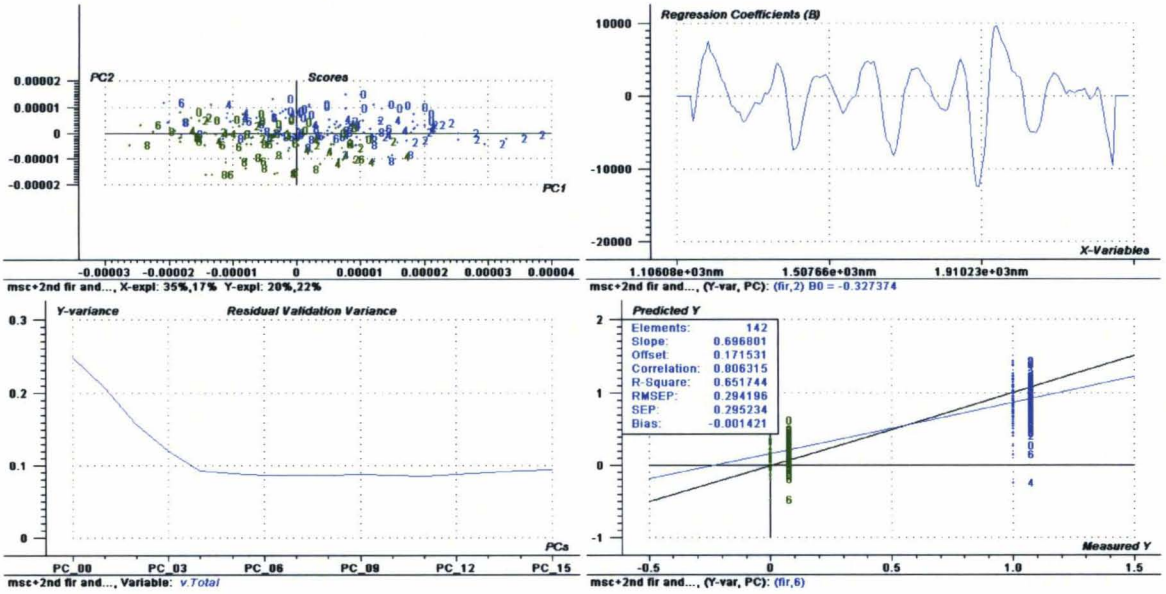


Fig. 3.61. PLS-DA regression validated model of fir and pine predictions whose surface was weathered for various periods: the upper left plot is Scores plot; the upper right plot is the Regression Coefficients plot for PC 1; the bottom left plot is the Residual Validation Variance plot; and the bottom right plot is the Predicted vs. Measured plot for fir with 6 PCs.



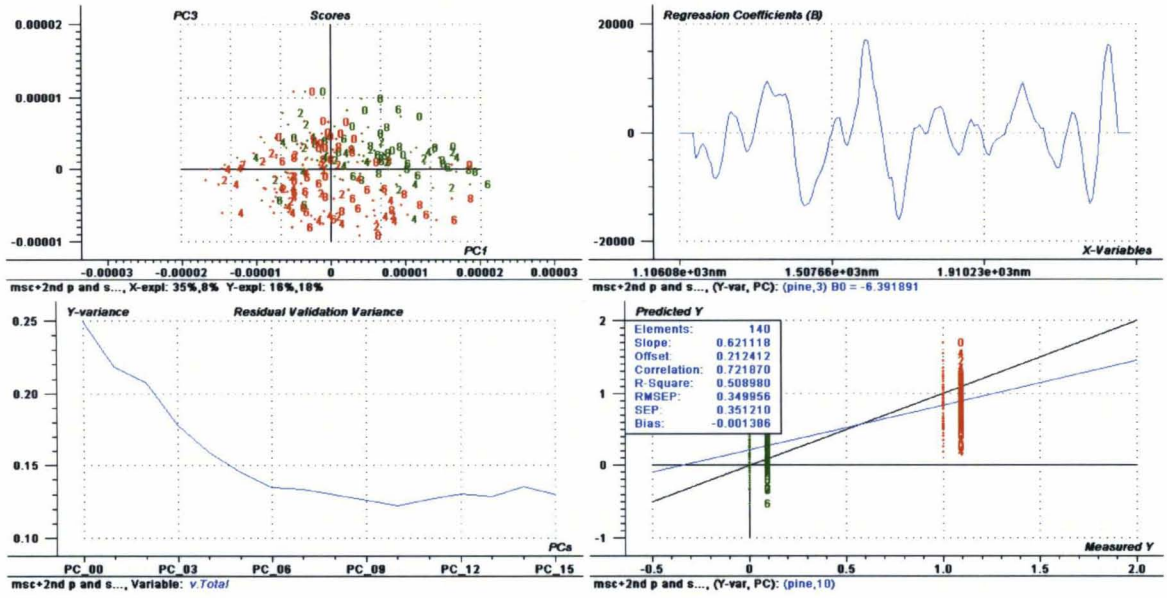
Red-pine; blue-fir; the regression coefficient plot for pine was the same but in opposite direction; and Predicted vs. Measured plot was the same for pine (not shown).

Fig. 3.62. PLS-DA regression validated model of fir and spruce predictions whose surface was weathered for various periods: the upper left plot is Scores plot; the upper right plot is the Regression Coefficients plot for PC 2; the bottom left plot is the Residual Validation Variance plot; and the bottom right plot is the Predicted vs. Measured plot for fir with 6 PCs.



Green-spruce; blue-fir; the regression coefficient plot for spruce was the same but in opposite direction; and Predicted vs. Measured plot was the same for spruce (not shown).

Fig. 3.63. PLS-DA regression validated model for pine and spruce predictions whose surface was weathered for various periods: the upper left plot is Scores plot; the upper right plot is the Regression Coefficients plot for PC 3; the bottom left plot is the Residual Validation Variance plot; and the bottom right plot is the Predicted vs. Measured plot for pine with 10 PCs.



Red-pine; green-spruce; the regression coefficient plot for spruce was the same but in opposite direction; and Predicted vs. Measured plot was the same for spruce (not shown).

Fig. 3.64. PLS-DA regression prediction model of all species whose surface was weathered for various periods: the upper left plot is Predicted vs. Reference plot of fir with 8 PCs; the upper right plot is Predicted vs. Reference plot of pine with 8 PCs; and the bottom plot is the Predicted vs. Reference plot of spruce with 8 PCs.

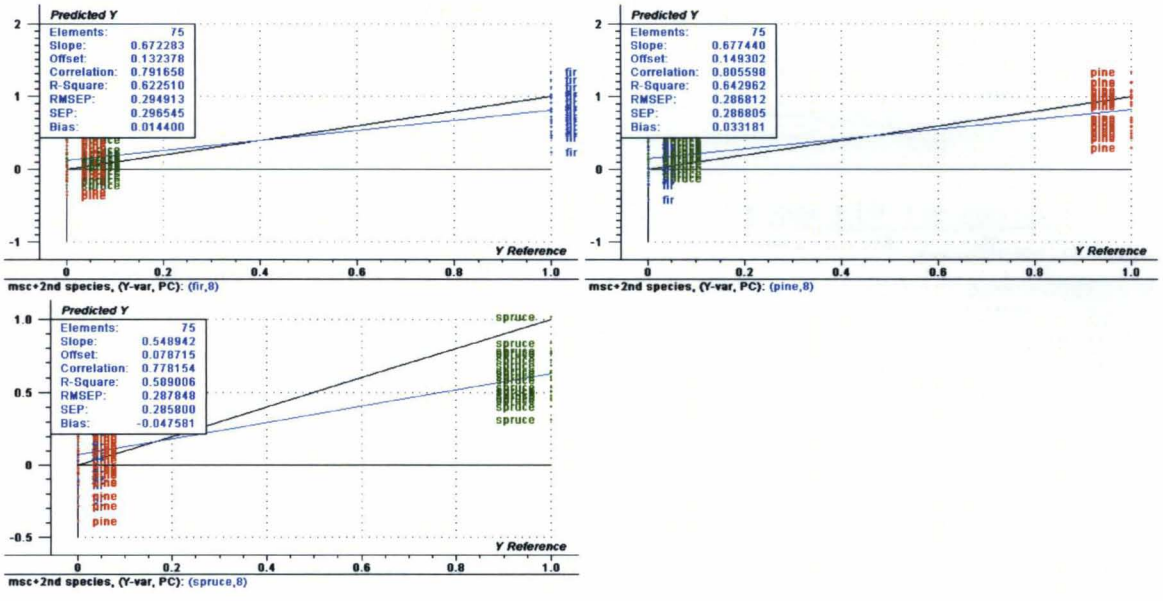


Fig. 3.65. PLS-DA regression prediction model of pine and fir whose surface was weathered for various periods: the upper plot is Predicted vs. Reference plot of pine with 6 PCs; and the bottom plot is the Predicted vs. Reference plot of fir with 6 PCs.

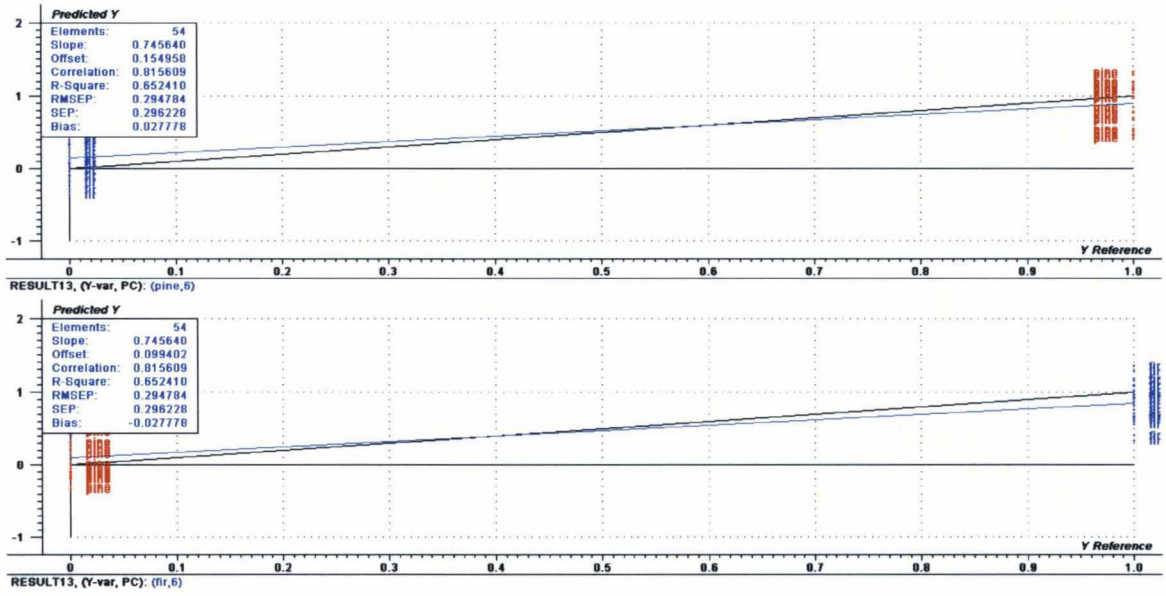


Fig. 3.66. PLS-DA regression prediction model of spruce and fir whose surface was weathered for various periods: the upper plot is Predicted vs. Reference plot of spruce with 6 PCs; and the bottom plot is the Predicted vs. Reference plot of fir with 6 PCs.

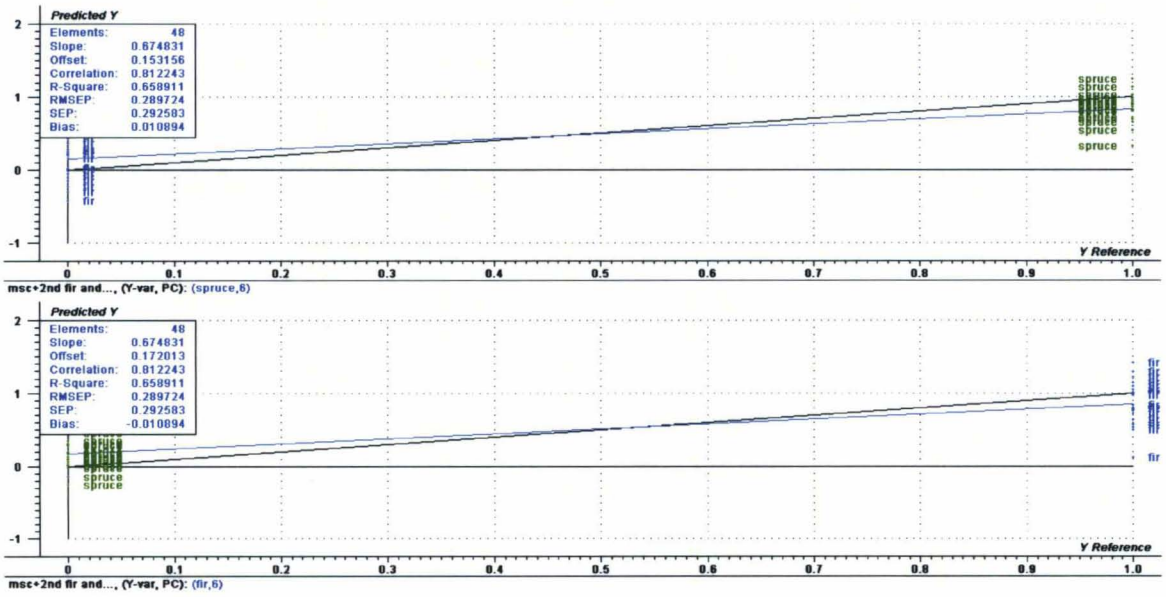
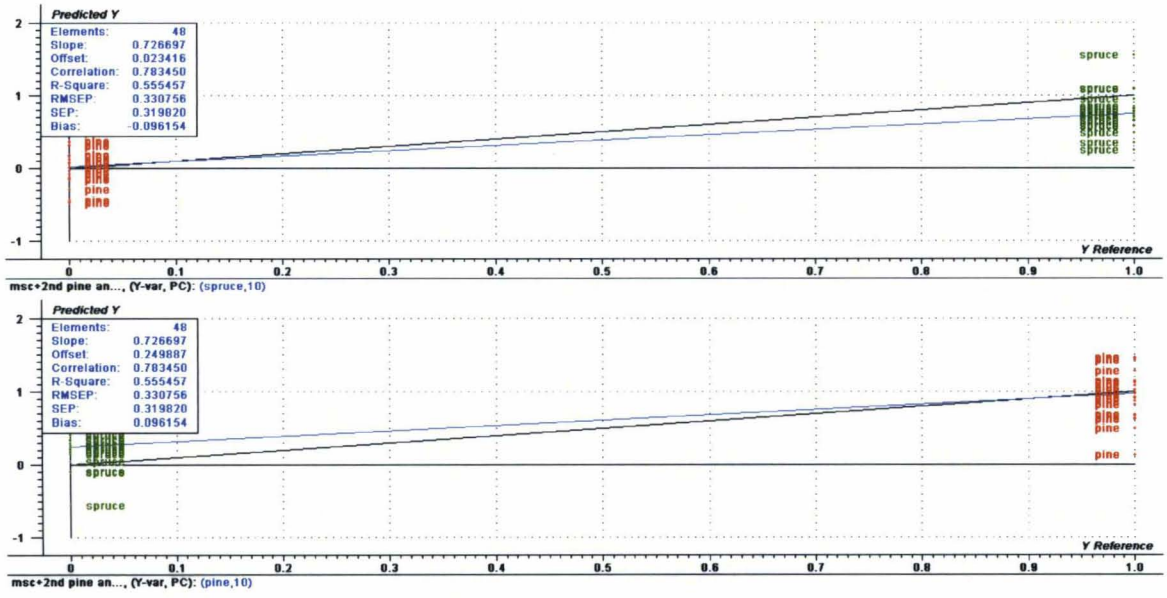


Fig. 3.67. PLS-DA regression prediction model of pine and spruce whose surface was weathered for various periods: the upper plot is Predicted vs. Reference plot of spruce with 10 PCs; and the bottom plot is the Predicted vs. Reference plot of pine with 10 PCs.



A validated model that simultaneously predicted species and weathering periods was developed (**Fig. 3.68**) and used to predict these variables. The calibration model had R^2 and RMSEP of 43% and 36%, respectively with 7 PCs for pine; 43% and 34% with 7 PCs for spruce; 56% and 32% with 7 PCs for fir; and 86% and 12% for weathering periods with 7 PCs. No significant outliers were detected. Prediction results for weathering periods and species are seen in **Fig. 3.69** and **Table 3.10**. The maximum error for species differentiation exhibited by the models predicting both species and weathering periods was 17% (13 out of 75 samples were misinterpreted) and therefore, both weathering periods and species could not be well predicted simultaneously when all the three species were included in the model. Better results were obtained when only two out of three species were used (**Fig. 3.70- 3.72** and **Table 3.10**).

Table 3.9. Prediction results for the validated PLS-DA regression models that predicted species whose surface was weathered for various periods.

Species included in the model	# of correct species type predictions (/27)		# of correct species type predictions (/21)	# PCs	R^2 (%)			RMSEP (/1)		
	Pine	Fir	Spruce		Pine	Spruce	Fir	Pine	Spruce	Fir
Spruce, fir, and pine	25*	26*	18*	8	64	59	62	0.29	0.29	0.29
Pine and fir	24**	25**	N/A	6	65	N/A	65	0.29	N/A	0.29
Fir and spruce	N/A	26***	20***	6	N/A	66	66	N/A	0.29	0.29
Pine and spruce	25****	N/A	18****	10	55	55	N/A	0.33	0.33	N/A

*1 pine was assigned fir and one pine was not assigned as any of the species; 1 fir was assigned as spruce; 1 spruce was assigned as fir and 1 spruce was not assigned as any of the species. **3 pines were assigned as fir; 2 firs were assigned as pine. ***1 fir was assigned as spruce and 1 spruce was assigned as fir. ****2 pines were assigned as pine; 2 spruces were assigned as pine and 1 spruce was not assigned to any of the species.

Fig. 3.68. PLS-DA regression validated model of both all species and weathering periods predictions: the upper left plot is Predicted vs. Measured plot for weathering periods predictions with 7 PCs; the upper right plot is the Predicted vs. Measured plot for pine with 7 PCs; the bottom left plot is the Predicted vs. Measured plot for spruce with 7 PCs; and the bottom right plot is the Predicted vs. Measured plot for fir with 7 PCs.

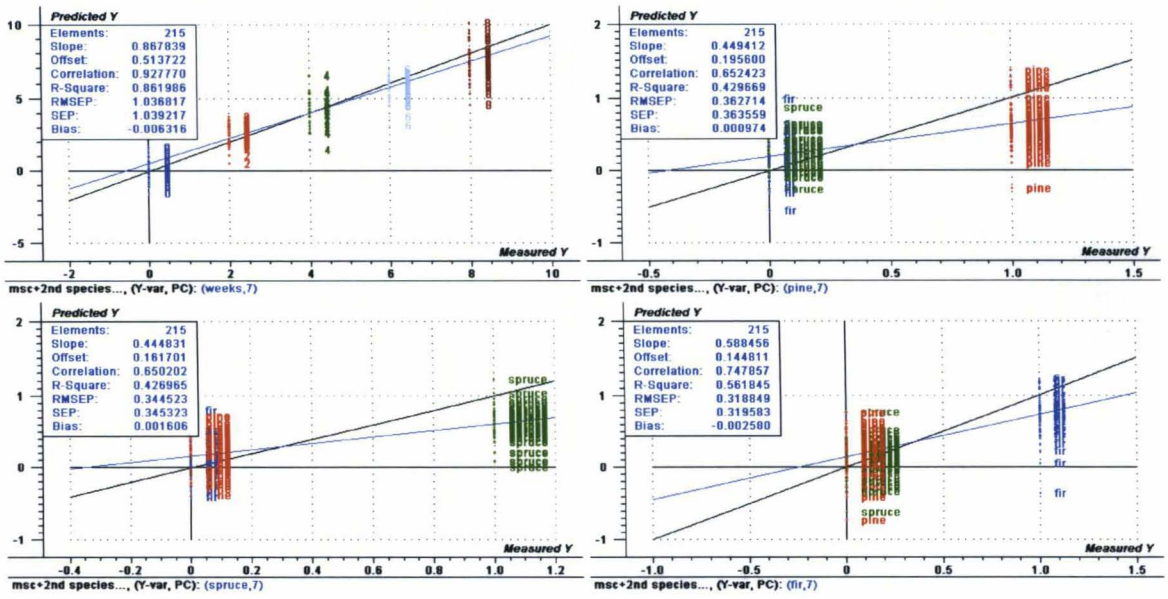


Fig. 3.69. PLS-DA regression prediction model of both all species and weathering periods: the upper left plot is Predicted vs. Reference plot for weathering periods prediction with 7 PCs; the upper right plot is the Predicted vs. Reference plot for fir with 7 PCs; the bottom left plot is the Predicted vs. Reference plot for pine with 7 PCs; and the bottom right plot is the Predicted vs. Reference plot for spruce with 7 PCs.

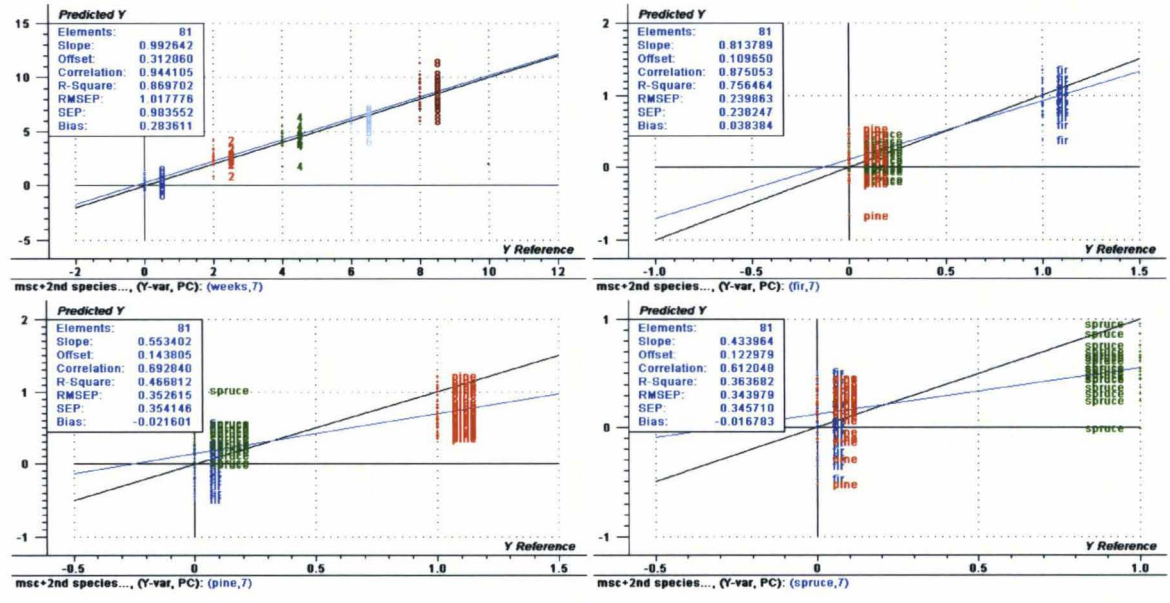


Fig. 3.70. PLS-DA regression prediction model of both pine and fir and weathering periods: the upper left plot is Predicted vs. Reference plot for weathering periods prediction with 3 PCs; the upper right plot is the Predicted vs. Reference plot for pine with 3 PCs; and the bottom plot is the Predicted vs. Reference plot for fir with 3 PCs.

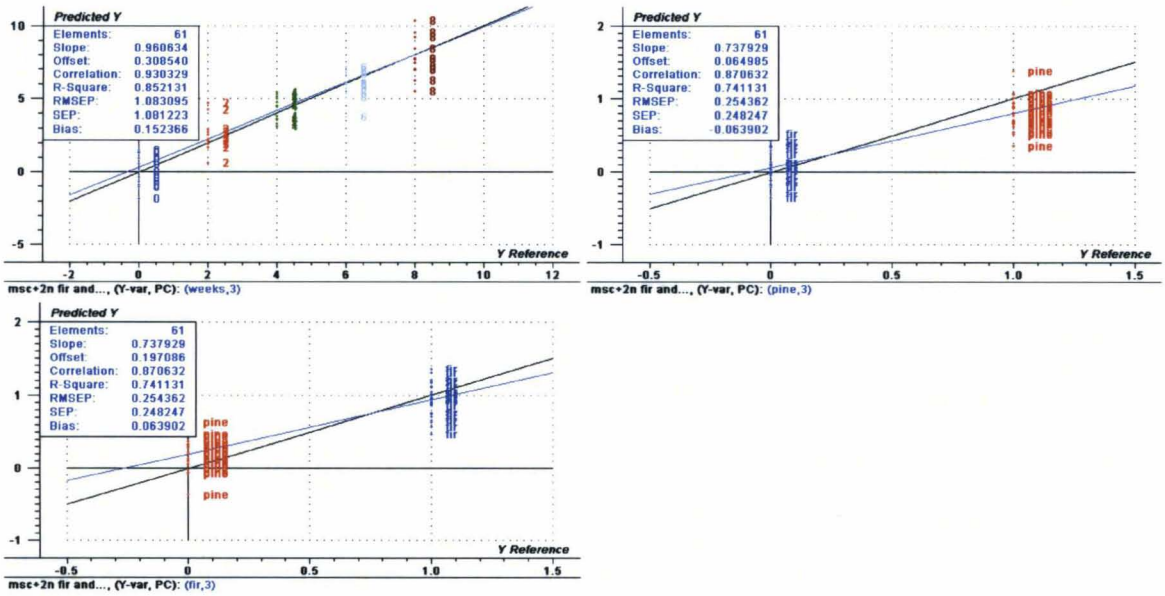


Fig. 3.71. PLS-DA regression prediction model of both spruce and fir and weathering periods: the upper left plot is Predicted vs. Reference plot for weathering periods prediction with 6 PCs; the upper right plot is the Predicted vs. Reference plot for spruce with 6 PCs; and the bottom left plot is the Predicted vs. Reference plot for fir with 6 PCs.

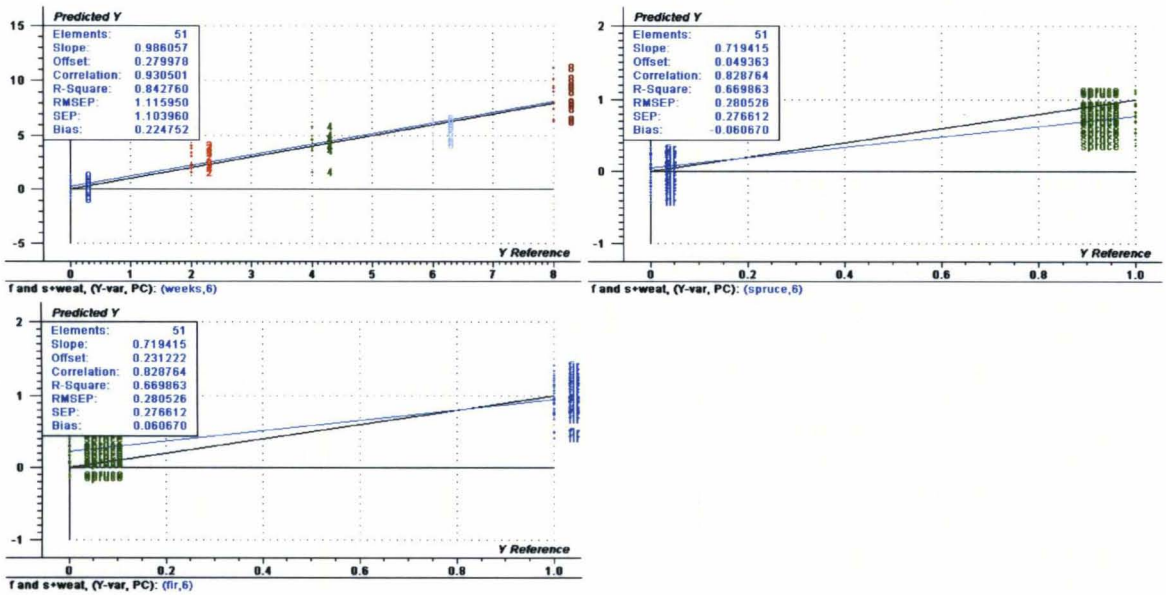
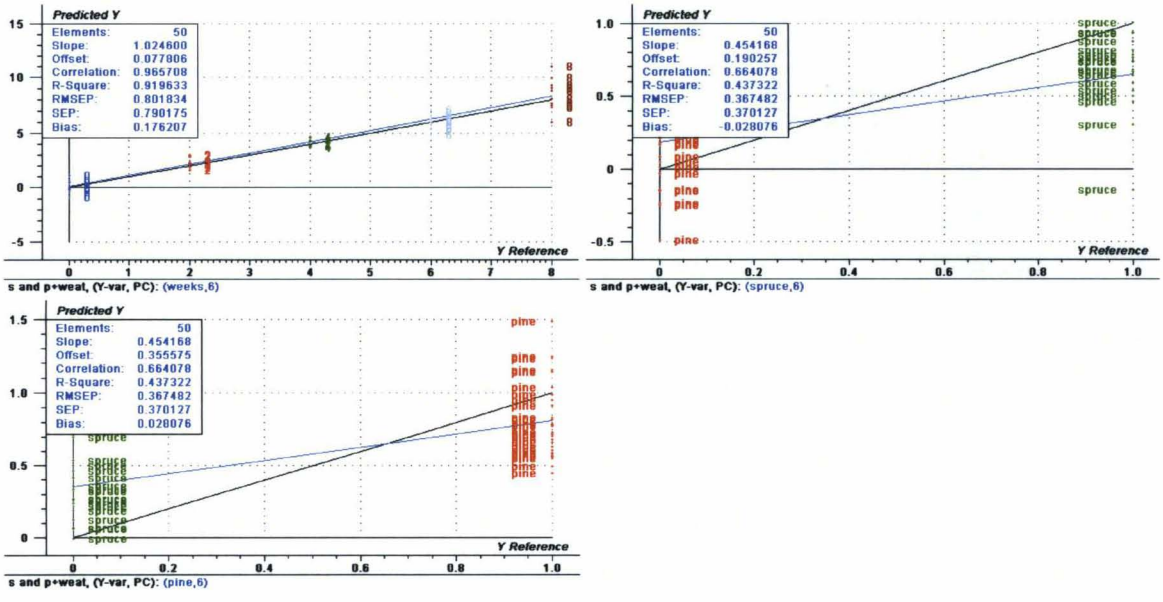


Fig. 3.72. PLS-DA regression prediction model of both spruce and pine and weathering periods: the upper left plot is Predicted vs. Reference plot for weathering periods prediction with 6 PCs; the upper right plot is the Predicted vs. Reference plot for spruce with 6 PCs; and the bottom left plot is the Predicted vs. Reference plot for pine with 6 PCs.



Finally, a calibration model consisting of non-weathered wood samples and samples weathered for six weeks was developed (**Fig. 3.73**) and used to predict samples weathered for six weeks with weathered surface removed or not removed. It was assumed that the weathered samples with the freshly cut surface should be similar to the non-weathered wood samples since weathering affected the wood surface only. Prediction results are shown in **Fig. 3.74**. 13 out of 15 weathered wood samples with freshly cut surfaces were assigned to non-weathered wood samples with R^2 and RMSEP of 69% and 28%, respectively; the rest were assigned to weathered wood samples whose surface was not cut (**Fig. 3.74**). In conclusion, the weathering did not affect the bulk of wood (at least to the depth of 10 mm) and so the confounding effect of weathering on species classification could be avoided by removing the weathered surface of wood to increase the prediction power.

Table 3.10. Prediction results for the validated PLS-DA calibration models that simultaneously predicted both species and weathering periods.

Species included in the model	# of correct species type predictions (/27)		# of correct species type predictions (/21)	# PCs	R^2 (%)				RMSEP (/1)			RMSEP (/8)
	Pine	Fir	Spruce		Weathering	Pine	Spruce	Fir	Pine	Spruce	Fir	Weathering
Spruce, fir, and pine	22*	25*	15*	7	87	47	36	76	0.35	0.34	0.24	1.02
Pine and fir	26**	26**	N/A	3	85	74	N/A	74	0.25	N/A	0.25	1.01
Fir and spruce	N/A	24***	19***	6	84	N/A	67	67	N/A	0.28	0.28	1.1
Pine and spruce	25****	N/A	17****	6	92	44	44	N/A	0.37	0.37	N/A	0.8

*1 pine was assigned as fir and 4 pines were not assigned to any of the species but closest to pine; 2 firs were not assigned to any of the species but closest to fir; 4 spruces were assigned as pine and 2 spruces were not assigned as any of the species but closest to spruce. **1 pine was assigned as fir and 1 fir was assigned as pine. ***3 firs were assigned as spruce; 1 spruce was assigned as fir and 1 spruce was not assigned to any of the species. ****2 pines were assigned as pine; 3 spruces were assigned as pine and 1 spruce was not assigned to any of the species.

Fig. 3.73. PLS-DA regression validated model of non-weathered wood samples and samples weathered for six weeks: the upper left plot is Scores plot; the upper right plot is the Regression Coefficients plot for PC 1; the bottom left plot is the Residual Validation Variance plot; and the bottom right plot is the Predicted vs. Measured plot for pine with 4 PCs.

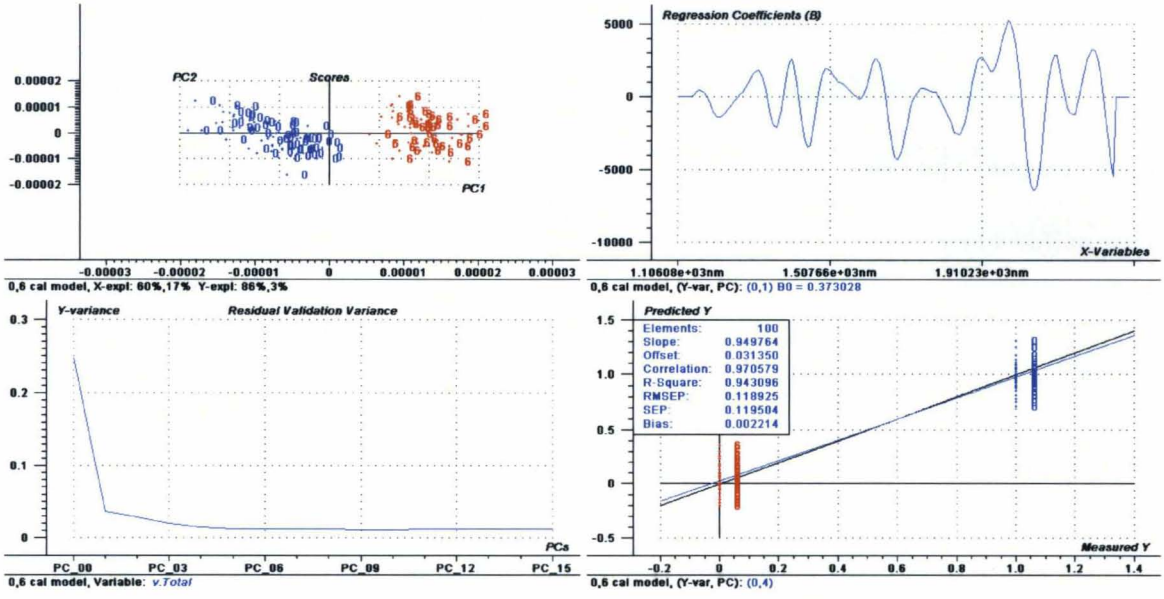
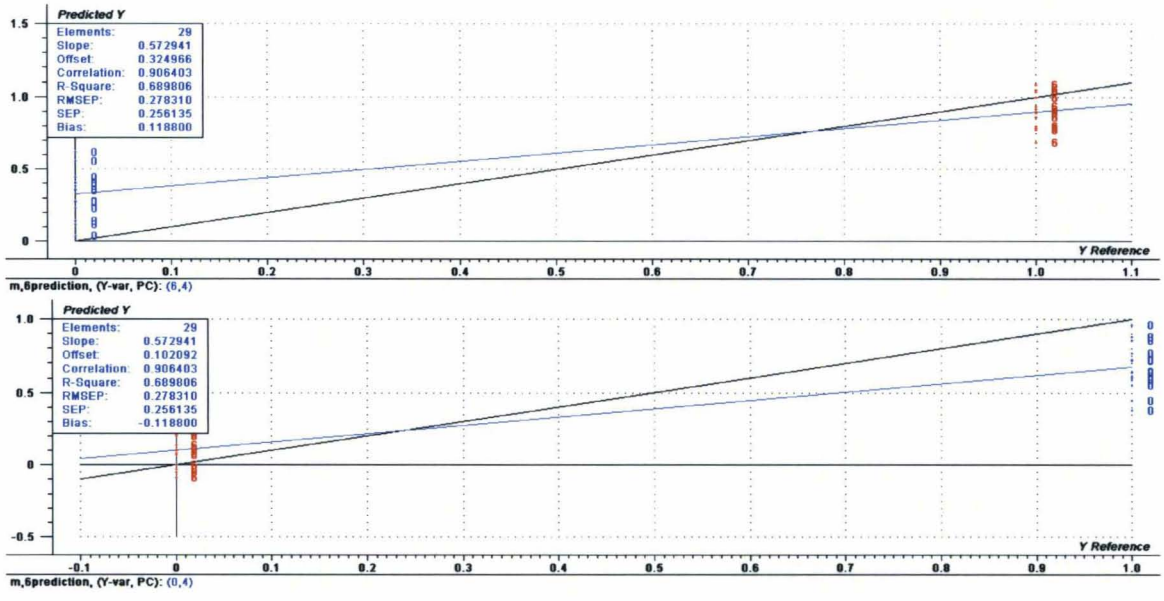


Fig. 3.74. PLS-DA regression prediction model of weathered wood samples with and without freshly cut surface: the upper plot is Predicted vs. Reference plot for wood samples weathered for six weeks with 4 PCs; and the bottom plot is the Predicted vs. Reference plot for wood samples weathered for six weeks having fresh surface with 4 PCs.



4. Conclusions

Optimal spectral treatment was determined. In this study the model with best predictive ability was obtained for normalized spectra transformed by MSC and second derivatization. The instrument reproducibility was also examined. The NIR spectrometer used in this study was found to be reproducible for at least two months.

Wood samples having constant MC and freshly cut surface were used to predict wood species. Separation of fir from both pine and spruce was superior to separation of pine and spruce. Luckily, reliably sorting fir wood species from SPF is more important. The maximum error for species differentiation exhibited by the models predicting species was 12%. Sorting by species is currently performed by eye and is difficult requiring an expert and always has an error associated with it. NIRS is amenable to automated analysis and has potential for quickly and relatively accurately identifying wood species. It would ease graders from a difficult task, while increasing product quality and decreasing production costs (51).

Since the wood EMC varies with the ambient RH and temperature the wood MC values could be well predicted (RMSEP of 6%) independent of wood species. All the current methods used for MC determination are either lengthy or have a high error associated with them especially for determination of MC above FSP. NIRS, on the other hand, is quick and has the potential to predict MC values more accurately. When wood samples conditioned to various MC values (oven-dry to FSP) were used for species

differentiation, the number of PCs and the prediction power increased slightly for each model compared to models where all wood species had a constant MC. The presence of the strong absorption bands in the NIR spectrum caused by moisture obscured spectral information derived from the wood and needed for species differentiation. It was not possible to delete the moisture wavelengths because the absorption at relevant wavelengths was so high that it impacted the whole NIR spectrum. However, species could still be predicted either separately or together with MC with a maximum error of 10%. Therefore, it is not required to condition all the wood samples to a constant MC saving analysis time.

Finally, since wood weathering leads to gradual surface degradation of wood and to eventual loss of lignin, the artificial weathering periods could be predicted independent of wood species. When wood samples at various weathering periods were used to predict species the models had lower prediction ability than the model composed of species having no weathering effect, indicating a confounding effect of weather. It was not possible to delete the weather-affected wavelengths because they contained information about chemical composition of the species. However, species could still be predicted either separately or together with weathering periods with a maximum error of 17%. If one wants to increase the prediction power, it is suggested to freshly cut the wood surface of all samples because weathering occurs only at the wood surface.

Further experiments with significantly larger sample sizes will be needed to confirm this level of accuracy. Improvements in the prediction models are also possible. The close

proximity of the probe to the wood in this study (2 mm) may not be appropriate for mill application of NIRS and light sources should be investigated that not only cover a larger area of wood, but that also allow NIR to be detected from a greater distance to the surface. Determining if inclusion of both heartwood and sapwood spectra in calibration models can improve prediction should also be investigated. Finally, since it is easier to analyze either tangential or radial wood planes on-line, additional NIR analysis should be performed on these planes.

5. References

1.Williams, P., & Norris, K. (Eds.). (1987). *Near-infrared technology in the agricultural and food industries.* St. Paul, Minnesota, USA: American Association of Cereal Chemists, Inc.

2. Burns, D. A., & Ciurczak, E. W. (2007). *Handbook of Near-Infrared Analysis* (3rd Ed.). Washington, D.C., USA: CRC Press.

3. Antal, I. & Dávid, Á. Z. (2007). *Fundamentals and Pharmaceutical Applications of Near-Infrared Spectroscopy.* Retrieved July 22, 2009, from http://www.glatt.com/ e/ 11_presse/times23site/tms23_fundamentals_1.htm

4. Esbensen, K., H. *Multivariate data analysis- In practice: An introduction to multivariate data analysis and experimental design* (5th Ed.). Woodbridge, NJ, USA: CAMO Smart Software Inc.

5. Rowell, R. M. (Ed.). (2005). *Handbook of Wood Chemistry and Wood Composites.* Washington, D.C., US: CRC Press.

6. Forest Products Laboratory. (1999). *Wood Handbook—Wood as an Engineering Material.* Madison, WI, US: Department of Agriculture, Forest Service, Forest Products Laboratory.

7. Axrup, L., Markides, K., & Nilsson, T. (2000). Using miniature diode array NIR spectrometers for analysing wood chips and bark samples in motion. *Journal of Chemometrics, 14*(5-6), 561-572.

8. Raymond, C. A., & Schimleck, L. R. (2002). Development of near infrared reflectance analysis calibrations for estimating genetic parameters for cellulose content in *Eucalyptus globulus. Canadian Journal of Forest Research, 32*(1), 170-176.

9. Poke, F. S., Wright, J. K., & Raymond, C. A. (2004). Predicting extractives and lignin contents in *Eucalyptus globulus* using near infrared reflectance analysis. *Journal of Wood Chemistry and Technology, 24*(1), 55-67.

10. Baillères, H., Davrieux, F., & Ham-Pichavant, F. (2002). Near infrared analysis as a tool for rapid screening of some major wood characteristics in a eucalyptus breeding program. *Annals of Forest Science, 59*(5-6), 479-490.

11. Brinkmann, K., Blaschke, L., & Polle, A. (2002). Comparison of different methods for lignin determination as a basis for calibration of near-infrared reflectance spectroscopy and implications of lignoproteins. *Journal of Chemical Ecology, 28*(12), 2483-2501.

12. Terdwongworakul, A., Punsuwan, V., Thanapase, W., & Tsuchikawa, S. (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. *Journal of Wood Science, 51*(2), 167-171.

13. Poke, F. S., & Raymond, C. A. (2006). Predicting extractives, lignin, and cellulose contents using near infrared spectroscopy on solid wood in eucalyptus globulus. *Journal of Wood Chemistry and Technology, 26*(2), 187-199.

14. Taylor, A. M., Freitag, C., Cadot, E., & Morrell, J. J. (2008). Potential of near infrared spectroscopy to assess hot-water-soluble extractive content and decay resistance of a tropical hardwood. *Holz Als Roh - Und Werkstoff, 66*(2), 107-111.

15. Zahri, S., Moubarik, A., Charrier, F., Chaix, G., Baillères, H., Nepveu, G., et al. (2008). Quantitative assessment of total phenol contents of European oak (*Quercus petraea* and *Quercus robur*) by diffuse reflectance NIR spectroscopy on solid wood

surfaces. *Holzforschung, 62*(6), 679-687.

16. Yeh, T. -.F, Chang, H. -.M, & Kadla, J. F. (2004). Rapid prediction of solid wood lignin content using transmittance near-infrared spectroscopy. *Journal of Agricultural and Food Chemistry, 52*(6), 1435-1439.

17. Sykes, R., Li, B., Hodge, G., Goldfarb, B., Kadla, J., & Chang, H. -M. (2005). Prediction of loblolly pine wood properties using transmittance near-infrared spectroscopy. *Canadian Journal of Forest Research, 35*(10), 2423-2431.

18. Kelley, S. S., Rials, T. G., Snell, R., Groom, L. H., & Sluiter, A. (2004). Use of near infrared spectroscopy to measure the chemical and mechanical properties of solid wood. *Wood Science and Technology, 38*(4), 257-276.

19. Yeh, T. -.F, Yamada, T., Capanema, E., Chang, H. -.M, Chiang, V., & Kadla, J. F. (2005). Rapid screening of wood chemical component variations using transmittance near-infrared spectroscopy. *Journal of Agricultural and Food Chemistry, 53*(9), 3328-3332.

20. Yamada, T., Yeh, T. -.F, Chang, H. -.M, Li, L., Kadla, J. F., & Chiang, V. L. (2006). Rapid analysis of transgenic trees using transmittance near-infrared spectroscopy (NIR). *Holzforschung, 60*(1), 24-28.

21. Chen, Q. -.M, Hu, Z., Chang, H. -.M, & Li, B. (2007). Micro analytical methods for determination of compression wood content in loblolly pine. *Journal of Wood Chemistry and Technology, 27*(3-4), 169-178.

22. Jones, P. D., Schimleck, L. R., Daniels, R. F., Clark III, A., & Purnell, R. C. (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*

Science and Technology, 42(5), 385-400.

23. Tsuchikawa, S., Takahashi, T., & Tsutsumi, S. (2000). Nondestructive measurement of wood properties by using near-infrared laser radiation. *Forest Products Journal, 50*(1), 81-86.

24. Jonsson, P., Sjöström, M., Wallbäcks, L., & Antti, H. (2004). Strategies for implementation and validation of on-line models for multivariate monitoring and control of wood chip properties. *Journal of Chemometrics, 18*(3-4), 203-207.

25. Defo, M., Taylor, A. M., & Bond, B. (2007). Determination of moisture content and density of fresh-sawn red oak lumber by near infrared spectroscopy. *Forest Products Journal, 57*(5), 68-72.

26. Adedipe, O. E., & Dawson-Andoh, B. (2008). Predicting moisture content of yellow-poplar (*Liriodendron tulipifera* L.) veneer using near infrared spectroscopy. *Forest Products Journal, 58*(4), 28-33.

27. Hoffmeyer, P., & Pedersen, J. G. (1995). Evaluation of density and strength of Norway spruce wood by near infrared reflectance spectroscopy. *Holz Als Roh- Und Werkstoff, 53*(1), 165-170.

28. Schimleck, L. R., Michell, A. J., Raymond, C. A., & Muneri, A. (1999). Estimation of basic density of *Eucalyptus globulus* using near-infrared spectroscopy. *Canadian Journal of Forest Research, 29*(2), 194-201.

29. Schimleck, L., R., Evans, R., & Ilic, J. (2001). Estimation of *Eucalyptus delegatensis* wood properties by near infrared spectroscopy. *Canadian Journal of Forest Research, 31*(10), 1671-1675.

30. Schimleck, L. R., Evans, R., & Matheson, A. C. (2002). Estimation of *Pinus radiata*

D. don clear wood properties by near-infrared spectroscopy. *Journal of Wood Science, 48*(2), 132-137.

31. Schimleck, L., Evans, R., & Ilic, J. (2001). Application of near infrared spectroscopy to a diverse range of species demonstrating wide density and stiffness variation. *IAWA Journal, 22*(4), 415-429.

32. Schimleck, L., Evans, R., & Ilic, J. (2003). Application of near infrared spectroscopy to the extracted wood of a diverse range of species. *IAWA Journal, 24*(4), 429-438.

33. Schimleck, L. R., Evans, R., Ilic, J., & Matheson, A. C. (2002). Estimation of wood stiffness of increment cores by near-infrared spectroscopy. *Canadian Journal of Forest Research, 32*(1), 129-135.

34. Schimleck, L. R., & Evans, R. (2002). Estimation of wood stiffness of increment cores by near infrared spectroscopy: The development and application of calibrations based on selected cores. *IAWA Journal, 23*(3), 217-224.

35. Schimleck, L. R., & Evans, R. (2002). Estimation of microfibril angle of increment cores by near infrared spectroscopy. *IAWA Journal, 23*(3), 225-234.

36. Schimleck, L. R., Mora, C., & Daniels, R. F. (2003). Estimation of the physical wood properties of green *Pinus taeda* radial samples by near infrared spectroscopy. *Canadian Journal of Forest Research, 33*(12), 2297-2305.

37. Kelley, S. S., Rials, T. G., Groom, L. R., & So, C. -L. (2004). Use of near infrared spectroscopy to predict the mechanical properties of six softwoods. *Holzforschung, 58*(3), 252-260.

38. Schimleck, L. R., Stürzenbecher, R., Mora, C., Jones, P. D., & Daniels, R. F. (2005). Comparison of *Pinus taeda* L. wood property calibrations based on NIR spectra from

the radial-longitudinal and radial-transverse faces of wooden strips. *Holzforschung, 59*(2), 214-218.

39. Jones, P. D., Schimleck, L. R., Peter, G. F., Daniels, R. F., & Clark III, A. (2005). Nondestructive estimation of *Pinus taeda* L. wood properties for samples from a wide range of sites in Georgia. *Canadian Journal of Forest Research, 35*(1), 85-92.
40. Schimleck, L. R., Jones, P. D., Clark III, A., Daniels, R. F., & Peter, G. F. (2005). Near infrared spectroscopy for the nondestructive estimation of clear wood properties of *Pinus taeda* L. from the southern united states. *Forest Products Journal, 55*(12), 21-28.
41. André, N., Labbé, N., Rials, T. G., & Kelley, S. S. (2006). Assessment of wood load condition by near infrared (NIR) spectroscopy. *Journal of Materials Science, 41*(7), 1879-1886.
42. Lestander, T. A., Lindeberg, J., Eriksson, D., & Bergsten, U. (2008). Prediction of *Pinus sylvestris* clear-wood properties using NIR spectroscopy and biorthogonal partial least squares regression. *Canadian Journal of Forest Research, 38*(7), 2052-2062.
43. Tsuchikawa, S., Hirashima, Y., Sasaki, Y., & Ando, K. (2005). Near-infrared spectroscopic study of the physical and mechanical properties of wood with meso- and micro-scale anatomical observation. *Applied Spectroscopy, 59*(1), 86-93.
44. Schimleck, L. R., Evans, R., Jones, P. D., Daniels, R. F., Peter, G. F., & Clark III, A. (2005). Estimation of microfibril angle and stiffness by near infrared spectroscopy using sample sets having limited wood density variation. *IAWA Journal, 26*(2), 175-187.
45. Fujimoto, T., Yamamoto, H., & Tsuchikawa, S. (2007). Estimation of wood stiffness

and strength properties of hybrid larch by near-infrared spectroscopy. *Applied Spectroscopy, 61*(8), 882-888.

46. Via, B. K., So, C. L., Shupe, T. F., Groom, L. H., & Wikaira, J. (2009). Mechanical response of longleaf pine to variation in microfibril angle, chemistry associated wavelengths, density, and radial position. *Composites Part A: Applied Science and Manufacturing, 40*(1), 60-66.

47. Hauksson, J. B., Bergqvist, G., Bergsten, U., Sjöström, M., & Edlund, U. (2001). Prediction of basic wood properties for Norway spruce. Interpretation of near infrared spectroscopy data using partial least squares regression. *Wood Science and Technology, 35*(6), 475-485.

48. Schimleck, L. R., Jones, P. D., Peter, G. F., Daniels, R. F., & Clark III, A. (2004). Nondestructive estimation of tracheid length from sections of radial wood strips by near infrared spectroscopy. *Holzforschung, 58*(4), 375-381.

49. Schimleck, L. R., & Evans, R. (2004). Estimation of *Pinus radiata* D. Don tracheid morphological characteristics by near infrared spectroscopy. *Holzforschung, 58*, 66-73.

50. Via, B. K., Shupe, T. F., Stine, M., So, C. L., & Groom, L. H. (2005). Tracheid length prediction in *Pinus palustris* by means of near infrared spectroscopy: The influence of age. *Holz Als Roh - Und Werkstoff, 63*(3), 231-236.

51. Brunner, C. C., Maristany, A. G., & Butler, D. A. (1996). Wood species identification using spectral reflectance. *Forest Products Journal, 46*(2), 82-85.

52. Antti, H., Sjöström, M., & Wallbäcks, L. (1996). Multivariate calibration models using NIR spectroscopy on pulp and paper industrial applications. *Journal of Chemometrics, 10*(5-6), 591-603.

53. Lebow, P. K., Brunner, C. C., Maristany, A. G., & Butler, D. A. (1996). Classification of wood surface features by spectral reflectance. *Wood Fiber Science 28*, 74-90.

54. Butler, D. A., Brunner, C. C., & Funck, J. W. (2001). Wood-surface feature classification using extended-color information. *Holz Als Roh - Und Werkstoff, 59*(6), 475-482.

55. Tsuchikawa, S., Inoue, K., Noma, J., & Hayashi, K. (2003). Application of near-infrared spectroscopy to wood discrimination. *Journal of Wood Science, 49*(1), 29-35.

56. Tsuchikawa, S., Yamato, K., & Inoue, K. (2003). Discriminant analysis of wood-based materials using near-infrared spectroscopy. *Journal of Wood Science, 49*(3), 275-280.

57. Flæte, P. O., Haartveit, E. Y., & Vadla, K. (2006). Near infrared spectroscopy with multivariate statistical modelling as a tool for differentiation of wood from tree species with similar appearance. *New Zealand Journal of Forestry Science, 36*(2-3), 382-392.

58. So, C. –H., Lebow, S. T., Groom, L. H., & Shupe, T. F. (2003). An evaluation of the use of near infrared (NIR) spectroscopy to identify water and oil-borne preservatives. In Managing the Treated Wood Resource- II: Special seminar sponsored by American Wood-Preservers' Association Utility Solid Waste Activities Group, May 1, Boston, MA.

59. So, C. -.L, Lebow, S. T., Groom, L. H., & Rials, T. G. (2004). The application of near infrared (NIR) spectroscopy to inorganic preservative-treated wood. *Wood and Fiber Science, 36*(3), 329-336.

60. Taylor, A., & Lloyd, J. (2007). Potential of near infrared spectroscopy to quantify

boron concentration in treated wood. *Forest Products Journal, 57*(1-2), 116-117.

61. Peydecastaing, J., Bras, J., Vaca-Garcia, C., Borredon, M. E., Iftimie, N., Giurginca, M., et al. (2006). NIR study of chemically modified cellulosic biopolymers. *Molecular Crystals and Liquid Crystals, 448,* 115-122.

62. Venås, T. M., & Rinnan, Å. (2008). Determination of weight percent gain in solid wood modified with in situ cured furfuryl alcohol by near-infrared reflectance spectroscopy. *Chemometrics and Intelligent Laboratory Systems, 92*(2), 125-130.

63. Kobori, H., Yonenobu, H., Noma, J., & Tsuchikawa, S. (2008). Discriminant analyzing system for wood wastes using a visible-near-infrared chemometric imaging technique. *Applied Spectroscopy, 62*(8), 854-859.

64. Schwanninger, M., Hinterstoisser, B., Gierlinger, N., Wimmer, R., & Hanger, J. (2004). Application of fourier transform near infrared spectroscopy (FT-NIR) to thermally modified wood. *Holz Als Roh - Und Werkstoff, 62*(6), 483-485.

65. Mitsui, K., Inagaki, T., & Tsuchikawa, S. (2008). Monitoring of hydroxyl groups in wood during heat treatment using NIR spectroscopy. *Biomacromolecules, 9*(1), 286-288.

66. Esteves, B., & Pereira, H. (2008). Quality assessment of heat-treated wood by NIR spectroscopy. *Holz Als Roh - Und Werkstoff, 66*(5), 323-332.

67. Kelley, S. S., Jellison, J., & Goodell, B. (2002). Use of NIR and pyrolysis-MBMS coupled with multivariate analysis for detecting the chemical changes associated with brown-rot biodegradation of spruce wood. *FEMS Microbiology Letters, 209*(1), 107-111.

68. Gierlinger, N., Jacques, D., Schwanninger, M., Wimmer, R., Hinterstoisser, B., &

Pâques, L. E. (2003). Rapid prediction of natural durability of larch heartwood using fourier transform near-infrared spectroscopy. *Canadian Journal of Forest Research, 33*(9), 1727-1736.

69. Flæte, P. O., & Haartveit, E. Y. (2004). Non-destructive prediction of decay resistance of *Pinus sylvestris* heartwood by near infrared spectroscopy. *Scandinavian Journal of Forest Research, Supplement, 19*(5), 55-63.

70. Ferraz, A., Mendonca, R., Guerra, A., Ruiz, J., Rodríguez, J., Baeza, J., et al. (2004). Near-infrared spectra and chemical characteristics of *Pinus taeda* (loblolly pine) wood chips biotreated by the white-rot fungus *Ceriporiopsis subvermispora*. *Journal of Wood Chemistry and Technology, 24*(2), 99-113.

71. Kent, S. M., Leichti, R. J., Morrell, J. J., Rosowsky, D. V., & Kelley, S. S. (2006). Analytical tools to predict changes in properties of oriented strandboard exposed to the fungus *Postia placenta*. *Holzforschung, 60*(3), 332-338.

72. Fackler, K., Gradinger, C., Hinterstoisser, B., Messner, K., & Schwanninger, M. (2006). Lignin degradation by white rot fungi on spruce wood shavings during short-time solid-state fermentations monitored by near infrared spectroscopy. *Enzyme and Microbial Technology, 39*(7), 1476-1483.

73. Sykacek, E., Gierlinger, N., Wimmer, R., & Schwanninger, M. (2006). Prediction of natural durability of commercial available European and Siberian larch by near-infrared spectroscopy. *Holzforschung, 60*(6), 643-647.

74. Fackler, K., Schwanninger, M., Gradinger, C., Hinterstoisser, B., & Messner, K. (2007). Qualitative and quantitative changes of beech wood degraded by wood-rotting basidiomycetes monitored by fourier transform infrared spectroscopic methods and

multivariate data analysis. *FEMS Microbiology Letters, 271*(2), 162-169.

75. Tsuchikawa, S., Inoue, K., & Mitsui, K. (2004). Spectroscopic monitoring of wood characteristics variation by light-irradiation. *Forest Products Journal, 54*(11), 71-76.

76. Wang, X., Wacker, J. P., & Rammer, D. R. (2006). Using NIR spectroscopy to predict weathered wood exposure times. In Proceedings WTCE- 9th World Conference on Timber Engineering, August 6-10, Portland, OR.

77. Yonenobu, H., & Tsuchikawa, S. (2003). Near-infrared spectroscopic comparison of antique and modern wood. *Applied Spectroscopy, 57*(11), 1451-1453.

78. Tsuchikawa, S., Yonenobu, H., & Siesler, H. W. (2005). Near-infrared spectroscopic observation of the ageing process in archaeological wood using a deuterium exchange method. *Analyst, 130*(3), 379-384.

79. Inagaki, T., Yonenobu, H., & Tsuchikawa, S. (2008). Near-infrared spectroscopic monitoring of the water adsorption/desorption process in modern and archaeological wood. *Applied Spectroscopy, 62*(8), 860-865.

80. Marklund, A., Paper, M., Hauksson, J. B., Edlund, U., & Sjöström, M. (1999). Prediction of strength parameters for softwood kraft pulps: Multivariate data analysis based on orthogonal signal correction and near infrared spectroscopy. *Nordic Pulp and Paper Research Journal, 14*(2), 140-148.

81. Fardim, P., Ferreira, M. M. C., & Durán, N. (2002). Multivariate calibration for quantitative analysis of eucalypt kraft pulp by NIR spectrometry. *Journal of Wood Chemistry and Technology, 22*(1), 67-81.

82. Alves, A., Santos, A., Da Silva Perez, D., Rodrigues, J., Pereira, H., Simões, R., et al. (2007). NIR PLSR model selection for kappa number prediction of maritime pine kraft

pulps. *Wood Science and Technology, 41*(6), 491-499.

83. Huang, A., Li, G., Fu, F., & Fei, B. (2008). Use of visible and near infrared spectroscopy to predict klason lignin content of bamboo, Chinese fir, paulownia, and poplar. *Journal of Wood Chemistry and Technology, 28*(3), 194-206.

84. Lindgren, T., & Edlund, U. (1998). Prediction of lignin content and pulp yield: From black liquor composition using near-infrared spectroscopy and partial least squares regression. *Nordic Pulp and Paper Research Journal, 13*(1), 76-80.

85. Raymond, C. A., Schimleck, L. R., Muneri, A., & Michell, A. J. (2001). Nondestructive sampling of *Eucalyptus globulus* and *E. nitens* for wood properties: III. predicted pulp yield using near infrared reflectance analysis. *Wood Science and Technology, 35*(3), 203-215.

86. Schimleck, L. R., Kube, P. D., Raymond, C. A., Michell, A. J., & French, J. (2006). Extending near infrared reflectance (NIR) pulp yield calibrations to new sites and species. *Journal of Wood Chemistry and Technology, 26*(4), 299-311.

87. Rowell, R. (Ed.). (1984). *Advances in Chemistry Series 207: The Chemistry of Solid Wood.* Washington, D.C., USA: American Chemical Society.

88. Hon, D.N.-S. (1981). Photochemical degradation of lignocellulosic materials. In: Grassie, N. (Ed.), *Developments in Polymer Degradation–3*. Essex: Applied Science Ltd. Chapter 8, 229–281.

89. Anderson, E.L., Pawlak, Z., Owen, N.L., & Feist, W.C. (1991). Infrared studies of wood weathering. Part I: Softwoods. *Applied Spectroscopy 45*(4):641–647.