DEPARTMENT OF CHEMISTRY, UNIVERSITY OF JYVÄSKYLÄ
RESEARCH REPORT No. 74

CHEMOMETRIC ADAPTATIONS IN WOOD PROCESSING CHEMISTRY

BY

PETTERI MALKAVAARA

Academic Dissertation
for the Degree of
Doctor of Philosophy

JYVÄSKYLÄN YLIOPISTO
Jyväskylä, Finland 2000
ISBN 951-39-0570-5
ISSN 0357-346X

DEPARTMENT OF CHEMISTRY, UNIVERSITY OF JYVÄSKYLÄ
RESEARCH REPORT No. 74

CHEMOMETRIC ADAPTATIONS IN WOOD PROCESSING CHEMISTRY

BY

PETTERI MALKAVAARA

Academic Dissertation
for the Degree of
Doctor of Philosophy

To be presented by the permission of the Faculty of Mathematics and Natural Sciences of the University of Jyväskylä for public examination in Auditorium S-212, on May 6th, 2000, at 12 o'clock noon

JYVÄSKYLÄN YLIOPISTO
Jyväskylä, Finland 2000
ISBN 951-39-0570-5
ISSN 0357-346X

URN:ISBN:978-952-86-0434-1
ISBN 978-952-86-0434-1 (PDF)
ISSN 0357-346X

University of Jyväskylä, 2024

The next room's light
that too goes out, and now –
the chill of night

Shiki (1867 - 1902)

Dedicated to my family, especially to my father, who somehow saw this coming (for so long time ago). You are remembered here, and always.

Muurame, 31st December, 1999

Petteri Malkavaara

LIST OF ORIGINAL PAPERS

This thesis is based on the following original publications. In the text they are referred to by their Roman numerals.

I Petteri Malkavaara, Pasi Harjula, Raimo Alén, and Juha Knuutinen, Chemometric Investigation on Structural Changes in Pine Kraft Lignin During Pulping, *Chemom. Intell. Lab. Syst.* accepted.
https://doi.org/10.1016/S0169-7439(00)00090-3

II Petteri Malkavaara, Raimo Alén, and Erkki Kolehmainen, Chemometrics: An Important Tool for the Modern Chemist; An Example from Wood Processing Chemistry, *J. Chem. Inf. Comp. Sci.* **40(2)** (2000) 438-441.
https://doi.org/10.1021/ci990444i

III Petteri Malkavaara, Raimo Alén, and Erkki Kolehmainen, Multivariate Correlation between ^{13}C and ^{31}P NMR Spectral Data on Dissolved Lignin and the Combustion Properties of Kraft Black Liquor, *Magn. Reson. Chem.* **37** (1999) 407-412.
https://doi.org/10.1002/(SICI)1097-458X(199906)37:6<407::AID-MRC475>3.0.CO;2-S

IV Petteri Malkavaara and Raimo Alén, A Spectroscopic Method for Determining Lignin Content of Softwood and Hardwood Kraft Pulps, *Chemom. Intell. Lab. Syst.* **44(1-2)** (1998) 287-292.
https://doi.org/10.1016/S0169-7439(98)00068-9

V Petteri Malkavaara and Raimo Alén, Diffuse Reflectance Vis and Chemometrics as a Useful Tool for Determining Pulp Yield and Kappa Number, *Tappi J.* accepted.
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2969368

VI Juha Hyötyläinen, Juha Knuutinen, Petteri Malkavaara, and Jussi Siltala, Pyrolysis-GC-MS and CuO-Oxidation-HPLC in the Characterization of HMMs from Sediments and Surface Waters Downstream of a Pulp Mill, *Chemosphere* **36(2)** (1998) 297-314.
https://doi.org/10.1016/S0045-6535(97)10006-6

VII Juha Hyötyläinen, Juha Knuutinen, and Petteri Malkavaara, Transport of High Molecular Mass Lignin Material in the Receiving Water System of a Mechanical Pulp Mill, *Chemosphere* **36(3)** (1998) 577-587.
https://doi.org/10.1016/S0045-6535(97)00366-4

PREFACE

This work has materialized at the Laboratory of Applied Chemistry, University of Jyväskylä during the last decade of a twentieth century. Financial support from Jenny and Antti Wihuri Foundation, Mikko Kaloinen Foundation, and the Academy of Finland is gratefully acknowledged.

I wish to express my deepest gratitude to my supervisor, Professor Raimo Alén for his support, encouragement and guidance during this work. I am expecially grateful to Professors Jorma Korvola, Juha Knuutinen, and Jussi Valkonen for their kind help and support. I owe my sincere thanks to all of my colleagues and co-workers, especially to Professor Erkki Kolehmainen (NMR-man!), Juha Hyötyläinen, Matti Ristolainen, Timo-Jaakko Toivanen, Marko Lehtonen, Pasi Harjula, Arja Mäkelä, Marja Salo, Kristiina Sakala, Reijo Kauppinen, Juhani Salovaara, Erkki Järvinen, Tapani Sorsa, and Veikko Manerus. I wish to extend my sincere thanks to the whole staff of the Department of Chemistry (especially to the laboratory of applied chemistry) for providing a pleasant and supporting atmosphere. It has truly been a privilege to work with you all.

I wish to express my special thanks to Docent Richard M. Ede (University of Waikato, New Zealand) and Professor Michael Sjöström (Umeå University, Sweden) for their valuable comments and suggestions regarding this thesis. I am especially indebted to Allan and Amanda Stewart (my dearest wood processors) for language revisions. Leif Riipinen (Oy Mitaten Finland Ab) is cordially thanked for numerous fruitful discussions.

Finally I wish to acknowledge my deepest gratitude to my family: Tiina, Mikko, Heikki and Juho. Your love, support and understanding has been and still is the most important achievement in my life: You are the shine in my eyes!

ABSTRACT

Multivariate chemometric data analysis techniques were applied to various research challenges mainly encountered in the selected process stages in the manufacture of kraft pulp. The applicability of these methods was demonstrated by several examples regarding characterization of the kraft lignin and determination of the combustion properties of kraft black liquors as well as determination of the kappa number, yield, and ISO brightness of kraft pulps. In addition, the structural elucidation of soluble high-molecular-mass lignin in receiving waters was performed.

In this study, various types of differently pre-treated spectroscopic and chromatographic data on the materials investigated were subjected to the principal component analysis (PCA), principal component regression (PCR), projection to latent structures (PLS) regression, and two-dimensional cross-correlation calculations (2DCC). Several types of predictive and descriptive models were established. These models functioned well within the experimental limits, producing results with good accuracy, and giving detailed information from inherently complex empirical datasets. Thus, it can generally be concluded that the adaptation of chemometric methods into wood processing applications will be of great benefit to the further development of these applications.

CONTENTS

ABBREVIATIONS

2CSAL	2-Chlorosyringaldehyde
2DCC	Two-dimensional cross-correlation
2D NMR	Two-dimensional nuclear magnetic resonance
4HACP	4-Hydroxyacetophenone
4HBAC	4-Hydroxybenzoic acid
4HBAL	4-Hydroxybenzaldehyde
4HCIN	4-Hydroxycinnamic acid
5FORV	5-Formylvanillin
6CACV	6-Chloroacetovanillone
6CVAC	6-Chlorovanillic acid
6CVAN	6-Chlorovanillin
ACSYR	Acetosyringone
ACVAN	Acetovanillone
BENAC	Benzoic acid
Bisphenol-A	2,2-bis(4-hydroxyphenol)propane
CIE	International Commission on Illumination
CP/MAS NMR	Cross-polarization/magic angle spinning solid state nuclear magnetic resonance
CZE	Capillary zone electrophoresis
Da	Dalton
DAD	Photodiode array detector
DDACV	Dehydrodiacetovanillone
DDVAN	Dehydrodivanillin
DMSO-d_6	Deuterated dimethylsulfoxide
FERAC	Ferulic acid
FID	Free induction decay
FTIR	Fourier transform infrared
GC	Gas chromatography
GC/MSD	Gas chromatography/mass selective detector

GC/NIRD	Gas chromatography/near infrared detector
GRAM	Generalized rank annihilation method
HPSEC	High performance size exclusion chromatography
ID	Inner diameter
IR	Infrared
ISO	International Organization for Standardization
LC	Liquid chromatography
LC/DAD	Liquid chromatography/photodiode array detector
L_{eff}	Effective length
L_{tot}	Total length
MLR	Multiple linear regression
MM	Molecular mass
M_n	Number-average molecular mass
M_w	Weight-average molecular mass
M_z	Z-average molecular mass
MPCA	Multiway principal component analysis
MRI	Magnetic resonance imaging
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MSD	Mass selective detector
NIPALS	Nonlinear iterative partial least squares
NIR	Near infrared
NMR	Nuclear magnetic resonance
N-PLS	Multilinear projection to latent structures
OSC	Orthogonal signal correction
PARAFAC	Parallel factor analysis
PC	Principal component
PCA	Principal component analysis
PCR	Principal component regression
PHTAC	Phthalic acid
PLS	Projection to latent structures

Py	Pyrolysis
R^2	Coefficient of determination
RAFA	Rank annihilation factor analysis
RSD	Relative standard deviation
S	Standard deviation
SCAN	Scandinavian Pulp, Paper and Board Testing Committee
SEC	Size exclusion chromatography
SEP	Standard error of prediction
SIMPLS	Simple way to do PLS
SVD	Singular value decomposition
SYALD	Syringaldehyde
SYRAC	Syringic acid
TAPPI	Technical Association of Pulp and Paper Industry
TLD	Tri-linear decomposition
TMS	Tetramethylsilane
UV	Ultraviolet
UV/Vis	Ultraviolet/visible light
VAN	Vanillin
VANAC	Vanillic acid

1 INTRODUCTION

Chemometrics can be defined as the chemical discipline that uses mathematical and statistical methods to relate measurements made on a chemical system to the state of the system and design or select optimal measurement procedures and experiments. Although a comprehensive definition of chemometrics is still a matter of debate, it is clearly seen that nearly all branches of chemistry would benefit by application of chemometric methods. For this reason, chemometrics has recently attracted increasing attention in chemical research and wood processing chemistry is no exception to that.

Wood processing chemistry uses analytical techniques adopted from analytical organic and inorganic chemistry. For monitoring processes and product quality there is, however, a clear need for rapid and accurate analytical methods and equipments. The proper integration of sophisticated spectroscopic and chromatographic determinations along with novel data handling techniques, such as multivariate analysis methods, is probably the most promising alternative for fulfilling this need.

In the year of 1998 160 million tons of wood pulp was produced and mainly used for manufacturing 298 million tons of paper and paperboard products.[1] The estimated value of paper and paperboard exports in 1998 is 70 000 million USD.[1] It is obvious that every aspect of the pulping and papermaking process is continuously reviewed very carefully to gain even the slightest improvement of the process. Also, understanding the fundamental chemistry which is the backbone of the processes, plays a critical role in the ongoing development in this research field.

2 DEFIBERING PROCESSES

The main basic product of the wood processing industry is pulp, predominantly used for papermaking.[2] In addition, some chemical pulps are processed to various cellulose derivatives, such as carboxymethylcellulose, cellulose acetate, and ethylhydroxyethylcellulose as well as regenerated celluloses (*e.g.*, rayon silk and cellophane). The main purpose of wood pulping is to liberate the fibers from the wood matrix, which can be accomplished either chemically or mechanically or by combining these two types of treatments. The common commercial pulps can be grouped into chemical, semichemical, chemimechanical, and mechanical types. The term high-yield-pulp is also often collectively used for different types of lignin-rich pulps which need mechanical defibration.

Chemical pulping (delignification) is a process in which lignin is removed from wood in the form that wood fibers are easily liberated either on the discharge from the digester or after a subsequent mild mechanical treatment.[2] Thus, the term delignification process refers to the selective removal of lignin. Practically all the production of chemical pulps in the world today is based on the sulfate (kraft) and sulfite processes, of which the former clearly predominates. Modern delignification processes can be divided into different phases which in most cases comprise modified kraft pulping, oxygen-alkali delignification, and bleaching.

2.1 KRAFT PULPING

Although today more than 80 % of the chemical pulp produced in the world is kraft pulp, there are still several drawbacks to the kraft process which are difficult to overcome.[2] The malodorous gases and the high consumption of bleaching chemicals in the case of softwood pulps are amongst these. However, according to the recent progress it is to be expected that new modifications will lead to improvements, not only to the product quality but also to the whole production process. The development of various modifications, such as pre-impregnation by steam or cooking liquors, liquor circulation systems, the design of cooking digester, and the use of additional chemicals (such as anthraquinone) have been common ways in facilitating the delignification.[3]

In kraft pulping the raw material, wood chips, are treated in either batch or continuous digesters at elevated pressures and temperatures (typically 1 - 2 hours in 160 - 180°C) with cooking liquor ("white liquor") consisting mainly of water, sodium hydroxide, and sodium hydrogen sulfide.[4] After the treatment pulp is removed from the reacted cooking liquor ("black liquor"), which contains the dissolved organic material along with the reacted and the unreacted cooking chemicals. Then pulp is washed, screened, and subjected to the further treatments depending on its end use. Black liquor is collected and burned in the recovery boiler[5] for energy production and regeneration of the cooking chemicals in the recovery cycle.[6] In addition, black liquor can be used as a raw material source (*e.g.*, the production of tall oil)[7] and also as a fill liquor at impregnation stages in the modified cooking systems.[8] Lately, other by-products of kraft process have been developed for commercial use.[9]

2.2 OXYGEN-ALKALI DELIGNIFICATION

In kraft pulping it is economically not feasible to dissolve all lignin present in wood feedstock since towards the end of cooking the cooking chemicals do not selectively react with lignin, leading to severe carbohydrate losses. To facilitate a more selective delignification, an oxygen-alkali delignification stage can be used prior to the actual bleaching sequence.[10] A typical process is carried out at 90 - 115°C under elevated pressure in alkaline conditions. Magnesium salts (*e.g.*, 0.1% on dry pulp) are also used to inhibit the effect of transition metals which otherwise would decompose the peroxides (formed in the process) to a harmful hydroxyl radicals. These radicals in turn would cause significant yield losses and excess consumption of delignification chemicals. About half of the residual lignin (*i.e.*, lignin present in pulp after cooking) can be removed using either single- or multistage oxygen-alkali delignification process with or without interstage washing, but a degree of lignin removal as high as 70% has been reported.[11]

2.3 BLEACHING

The purpose of bleaching is to increase the brightness and purity of pulp. This can be done by either selectively removing or brightening the color-causing functional groups (chromophores), mainly originating from lignin. Typical bleaching agents are chlorine dioxide, hypochlorite, hydrogen peroxide, and ozone.[12] Other bleaching stages are, *e.g.*, selective acid treatment, alkaline extraction, chelation, and the use of enzymes (mainly xylanases and glucomannanases). In Finland the use of chlorine gas as a bleaching agent ended in 1994.

A general practise in the bleaching of chemical pulp is to alternate acid and alkaline stages.[13] In the acid stage an oxidative agent is applied, and in the subsequent alkaline stage the oxidation products are extracted from the pulp. Exceptions to this principle are hypochlorite and peroxide stages, where the oxidation are carried out in the alkaline medium. With these agents the oxidative bleaching and alkaline extraction can be performed in the same stage. In addition, the bleaching of chemical pulps decreases the amount of residual extractives.[12]

3 MULTIVARIATE DATA ANALYSIS

3.1 PRINCIPAL COMPONENT ANALYSIS

Closely related to factor analysis, principal component analysis (PCA) is the basis for a multitude of multivariate analysis techniques. In the matrix form, PCA can be expressed as

$$\mathbf{X} = \mathbf{TP'} + \mathbf{E}$$

where **X** usually describes the variable mean centered and scaled original data, **T** is a so-called score matrix, **P'** is a loading matrix, and **E** describes the residual error (*i.e.*, noise).[14,15] The number of principal components (rank) is the number of columns in **T** and rows in **P'** matrices. All the principal components (PCs) are orthogonal to each other and they represent successively smaller and smaller variances of the original data, *i.e.*, the first PC consists of the largest variance in the original data (**X**) and the last PC consists of the smallest variance. The loading and score matrices are calculated using either a singular value decomposition (SVD) or NIPALS (Nonlinear Iterative Partial Least Squares) algorithm. The difference between those algorithms is that NIPALS calculates PCs successively one by one (meaning also that the number of PCs to be calculated can be chosen), whereas with SVD, all possible PCs are calculated at the same time. Other decomposition methods, such as SIMPLS algorithm, are also proposed for eigenvalue calculations. [16]

In this type of analysis, the score values represent the projections of the original objects (*i.e.*, samples) in a given principal coordinate system defined by the loading vectors calculated from the original variables. If the variables are, for example, NMR spectral data, then each loading vector can be plotted against an original NMR scale (in the form of a "loading subspectrum") and thus interpreted like a NMR spectrum. In addition, the resonances with the highest loading values mean in this case that there are systematic changes between the original spectra in these areas. Samples obtaining high absolute score values in that principal component can be concluded as having the largest amount of systematic variation in the variables with high loadings.

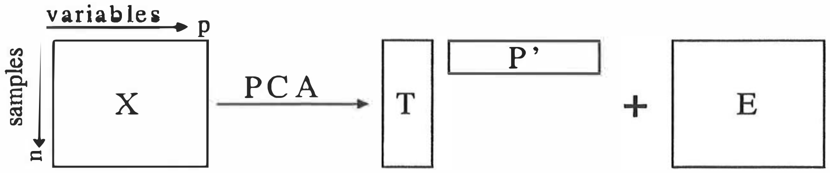

Figure 1. Schematic presentation of principal component analysis. For explanation of symbols, see text.

PCA can also be regarded as a tool for decreasing the dimensionality of a multidimensional system without losing any essential information. For example, in drug design, the molecular design techniques in combinatorial chemistry can be used to provide as large a molecular diversity as possible for the combinatorial library, which is to be tested for the desired biological activity. This molecular diversity is sought by plotting the relevant molecular properties against each other to obtain projections of molecules in the multidimensional space (which may involve as many as 20 or 30 dimensions). PCA can efficiently be used to decrease the dimensionality of the analysis by creating a low-dimensional model from the few significant principal components, thus facilitating decision-making.[17]

PCA can also be applied as a noise reduction technique. When the original data set is divided into its principal components, it can be reconstructed by taking only the most meaningful principal components (*i.e.*, those having the largest proportion of the modelled variance) into account. The result would then be the noise-reduced original data.

3.2 PRINCIPAL COMPONENT REGRESSION

Principal component regression (PCR) is a regression method for multivariate data. In contrast to the multiple linear regression (MLR) method, PCR does not place restrictions to the correlation amongst the independent (explaining) variables. Instead of using original, possible inter-correlating explaining variables (**X**), their score values, obtained from PCA, are used and the dependent variable or variables (**Y**) are regressed on them. These score values are by definition orthogonal and therefore not inter-correlating, thus facilitating greatly the regression analysis.

3.3 PROJECTION TO LATENT STRUCTURES

Projection to latent structures (sometimes referred to as partial least squares) regression (PLS) is similar to PCR but has marked differences: a decomposition is performed also to the dependent variable or variables. The variance ("structure") from **Y** is also taken account in calculating the regression coefficients. This is accomplished by a so-called "inner relationship" which maximizes the covariance between the **X** and **Y** scores. In PLS, the calculated factors are called latent variables, and are numerically different from principal components of **X** and **Y**.

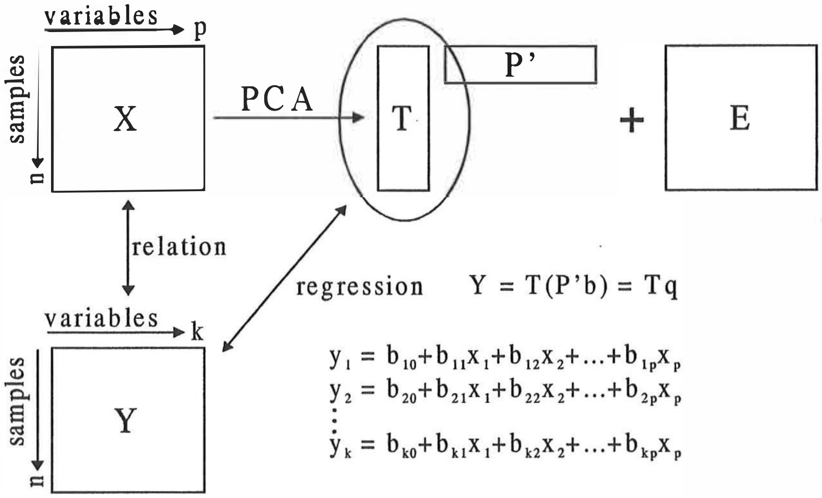

Figure 2. Schematic presentation of principal component regression.

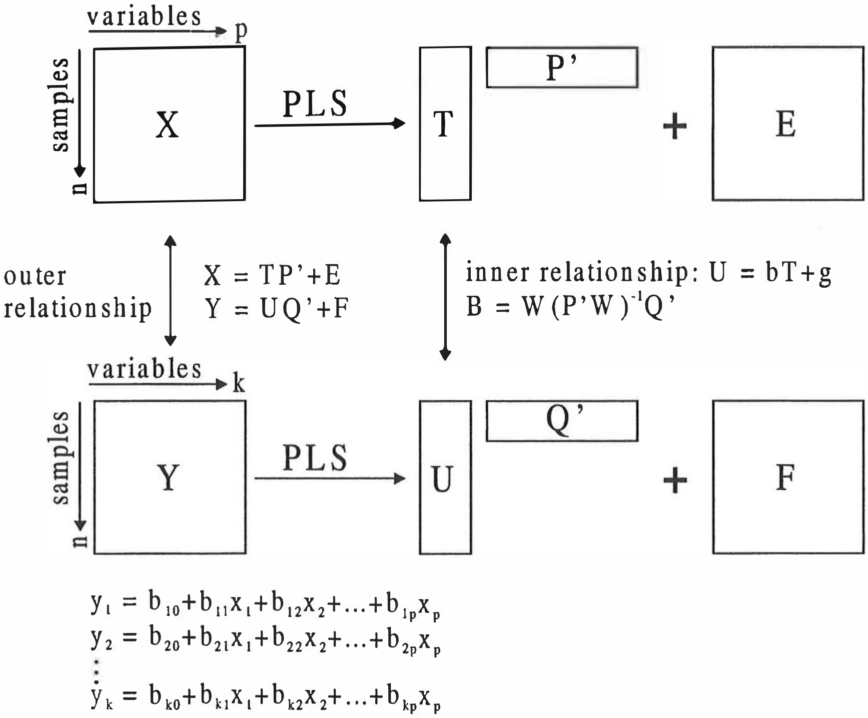

Figure 3. Schematic presentation of projection to latent structures regression.

3.4 TWO-DIMENSIONAL CROSS-CORRELATION CALCULATIONS

Two-dimensional cross-correlation (2DCC) calculations are a more general approach to a multidimensional spectroscopy.[18] The idea is simply to apply statistical cross-correlation calculations, based on the chemometric correlation calculations in one dimension, to the higher-dimensional data. The results from these 2DCC calculations would highlight the simultaneous changes in the correlated data and offer a unique tool to assign correlations to the spectral regions. In addition, the statistical cross-correlation approach is not limited only to the spectroscopy: any relevant analysis data, regardless of its type, can be subjected to the calculations. In the case of point-for-point 2DCC calculations, the correlation coefficients (r_{ij}) are calculated according to equation (1):[18]

$$\boldsymbol{R}_{i,j} = \frac{n\sum x_j y_i - \left(\sum x_j\right)\left(\sum y_i\right)}{\sqrt{\left[n\sum x_j^2 - \left(\sum x_j\right)^2\right]\left[n\sum y_i^2 - \left(\sum y_i\right)^2\right]}} \qquad \text{Eqn (1)}$$

where **R** is the matrix consisting of correlation coefficients r_{ij}, x_j represent a vector in the first spectroscopic data set, and y_i in the second.

3.5 ORTHOGONAL SIGNAL CORRECTION

A variant of PLS used to achieve a signal correction that is as close to orthogonal as possible to a given y vector or **Y** matrix is called orthogonal signal correction (OSC). In the method the variation in **X** which is not correlated, *i.e.*, is orthogonal to **Y** is first maximized and then subtracted from the original **X**. The new, OSC-corrected **X**, denoted (**X***) is used to subsequent PLS modelling.[19]

3.6 MULTIWAY MODELS

PCA, PCR, and PLS are bilinear projection methods (two-way methods) typically used in the multivariate analysis of data obtained by first order instrument, *i.e.*, analytical instrument which produces a vector of data per sample (spectroscopy, liquid chromatography (LC), gas chromatography (GC), and sensor arrays). However, many advanced analytical instruments are of second order, producing a matrix of data per sample. Some examples of these are excitation emission fluorescence, two-dimensional nuclear magnetic resonance (2D NMR), tandem mass spectrometry (MS/MS), magnetic resonance imaging (MRI), and "hyphenated instruments" like GC/MSD, GC/NIRD, and LC/DAD.[20]

Analytical data of second (or higher) order can be subjected to a multiway (multi, meaning three or higher) modelling using methods like rank annihilation factor analysis (RAFA) and generalized rank annihilation method (GRAM)[21-23], tri-linear decomposition (TLD)[24], parallel factor analysis (PARAFAC)[25], multiway PCA (MPCA)[26],Tucker3 method[27], and multilinear PLS regression (N-PLS)[28]. Three-way models are considered as three-dimensional extension of two-way models, *e.g.*, PARAFAC and Tucker3 are multilinear analogues of PCA, N-PLS being a multilinear analogue of PLS.

Multivariate image analysis with PCA is an exception to true multiway analysis and is considered as a weak multiway analysis technique. Image PCA is an unfolding PCA method which does not always give physically meaningful decomposition, but is still useful in certain applications.[29]

3.7 APPLICATIONS IN WOOD PROCESSING CHEMISTRY

Multivariate analysis techniques have already in some applications proved to be useful tools for the evaluation of spectral and chemical data on different pulps.[30-42] In the previous investigations[43-48], the lignin content and the carbohydrate composition of both intact wood and pulp samples were obtained by applying multivariate analysis techniques (PCA, PCR, and PLS) to various spectral data obtained mainly on the basis of infrared (IR) and cross-polarization/magic angle spinning solid state ^{13}C nuclear magnetic resonance (CP/MAS ^{13}C NMR) determinations.

The rapid determination of the chemical composition and density of *Pinus radiata* has been accomplished by PLS modelling of FTIR data[49], whereas PLS of free induction decay (FID) of NMR pulse was used to determine the moisture content and basic density of softwood above fiber saturation.[50] The moisture content and basic density of Norway spruce (*Picea abies*) have also been determined using near infrared (NIR) spectroscopy with multivariate data analysis.[51]

The storage time of wet-stored timber has been estimated along with the sapwood and heartwood classification by multivariate analysis of NIR spectral data on logs of Scots pine (*Pinus sylvestris*).[52] Another study indicated that it may be possible to differentiate the samples of various wood species and samples of different origins using similar techniques.[53] The compression strength and density along with chemical and biological degradation of Norway spruce (*Picea abies*) has been predicted using multivariate data analysis and NIR.[54] Paper and pulp strength parameters have been determined spectroscopically using both mid and near FTIR data combined with chemometrics.[55]

NIR spectroscopy, OSC, and PLS were used to determine some physical properties of laboratory-cooked kraft pulps[19], and the prediction of strength parameters for softwood kraft pulp using NIR and PLS either with or without OSC have been recently reported.[56,57]

Alkali profiles[58] along with lignin content and pulp yield[59] have been modelled by multivariate analysis of NIR spectral data on kraft black liquors. Within-tree variations of pulp yield in plantation-grown trees of *Eucalyptus nitens* were modelled by multivariate analysis of NIR spectral data.[60] The hydroxymethylation of acetosolv lignin from sugarcane bagasse was studied by means of FTIR and PCA.[61]

Fluorescence spectra of pulps[62] and (in conjunction with ^{31}P NMR) lignin[63] have been studied using multivariate data analysis techniques, and pulping process liquors have been analyzed by UV/Vis spectroscopy and chemometrics.[64] Heterogeneous photocatalysis treatment of kraft process effluents[65] and combustion properties of black liquors[66] have been studied using multivariate analysis.

The dependence of measurement error on sampling frequency and the role of chlorite consentration in predicting brightness at D1 chlorine bleaching stage have been studied using chemometrics.[67,68] Wastewater treatment of pulp and paper mill and monitoring the papermaking process data have been studied using advanced multivariate analytical methods.[69-74]

3.8 AIM OF THIS WORK

The main goal of this thesis was to apply various chemometric data analysis techniques to several analytical research examples (cases) in the area of wood processing chemistry. These cases were consciously selected to represent typical stages or challenging analytical points of the wood processing chain "from wood to paper". However, the main emphasis was paid to chemical defibering processes based on kraft pulping.

4 EXPERIMENTAL

4.1 KRAFT PULPING AND OXYGEN-ALKALI DELIGNIFICATION

Laboratory-scale pulping experiments[I] on screened industrial chips of Scots pine (*Pinus sylvestris*) were carried out in a battery of stainless steel autoclaves, each charged with 130 g of oven-dry chips. The autoclaves were rotated in a decene bath under the conditions shown in Table 1. The total cooking yield, kappa number (according to the TAPPI T 236 cm-85 standard), and ISO brightness (according to the ISO 2470 standard) values of the corresponding pulp samples are also listed in Table 1.

Table 1. Cooking data* and total cooking yield, kappa number, and ISO brightness values of pulp samples[I]

Sample	SW1	SW2	SW3	SW4	SW5	SW6	SW7	SW8
Time at max T (min)	10	40	70	100	130	160	190	210
Total cooking yield	56,6	43,7	41,3	40,4	37,6	37,1	36,7	35,7
Kappa number**	72,8	23,7	16,9	13,7	10,6	10,1	9,2	8,3
ISO brightness***	26,4	32,5	32,9	33,7	36,8	36,1	36,2	37,0

* Heating-up time (from 85°C to 175°C): 90 min, liquor-to-wood ratio: 5.0 L/kg, alkali charge: 35 % active alkali (as NaOH) on wood, sulfidity 30 %.
** TAPPI T 236 cm-85 standard.
*** ISO 2470 standard.

Laboratory-scale pulping experiments[II] on screened industrial chips of Scots pine (*Pinus sylvestris*) and silver birch (*Betula pendula*) were carried out in a battery of stainless steel autoclaves, each charged with 130 g of oven-dry chips. The autoclaves were rotated in a decene bath under the conditions shown in Table 2.

Table 2. Cooking data[II]

Pine kraft pulping*								
Time at max T (min)	10	40	70	100	130	160	190	210
Sample	SW1	SW2	SW3	SW4	SW5	SW6	SW7	SW8
Birch kraft pulping**								
Time at max T (min)	0	5	10	20	30	40	50	60
Sample	HW1	HW2	HW3	HW4	HW5	HW6	HW7	HW8
Time at max T (min)	80	100	140	200				
Sample	HW9	HW10	HW11	HW12				

*Heating-up time (from 85°C to 175°C): 90 min
Liquor-to-wood ratio: 5.0 L/kg
Alkali charge: 35 % active alkali (as NaOH) on wood, sulfidity 30 %.
**Heating-up time (from 80°C to 170°C): 90 min
Liquor-to-wood ratio: 4.5 L/kg
Alkali charge: 23.2 % active alkali (as NaOH) on wood, sulfidity 33.4 %.

Air-dried industrial chips of pine wood (*Pinus sylvestris*) and birch wood (*Betula pendula*) were delignified in a laboratory-scale digester under the conditions shown in Table 3, which also indicates the calculated amount of dissolved lignin in each case.[III]

Table 3. Total cooking yield and amount of dissolved lignin as a function of cooking time[III]*

Sample	Cooking time, min	Cooking yield, %	Dissolved lignin, % of the initial**	Sample	Cooking time, min	Cooking yield, %	Dissolved lignin, % of the initial**
Pine : birch (100 : 0)				**Pine : birch (60 : 40)**			
A1	115	55,6	51,4	**C1**	109	54,8	52,6
A2	140	50,5	71,1	**C2**	128	50,5	73,5
A3	165	48,6	75,2	**C3**	147	49,1	78,0
A4	190	47,2	81,2	**C4**	166	49,2	80,9
A5	215	46,4	78,1	**C5**	185	48,4	83,9
Pine : birch (80 : 20)				**Pine : birch (0 : 100)**			
B1	109	56,3	46,1	**D1**	105	54,8	65,8
B2	128	51,6	65,1	**D2**	120	51,9	82,4
B3	147	49,8	69,3	**D3**	135	51,7	82,6
B4	166	48,8	75,9	**D4**	150	51,9	84,6
B5	185	47,4	81,8	**D5**	165	51,2	82,4

* The cooking conditions: Active alkali 20 % (NaOH) on wood, sulfidity 30 %, liquor-to-wood ratio 4.0 L/kg, heating-up time (20 to 170°C) 90 min.

** Calculated values based on the initial amount of lignin in the wood feedstock and in the corresponding pulps (obtained from the chlorine number determination, SCAN-C 29:72).

Two sets of unbleached kraft pulp samples with varying kappa numbers (Table 4) were produced by laboratory-scale kraft cooking of Scots pine (*Pinus sylvestris*), Norway spruce *(Picea abies)*, and silver birch (*Betula verrucosa*).[IV] Set I, conventional cooking, consisted of 65 pine, spruce, and birch pulp samples, and Set II, displacement cooking, consisted of 56 pine and birch pulp samples. In addition, selected samples were further delignified (Sets III, IV, and V) using oxygen under the following conditions: consistency 5 %, NaOH charge 5 % on pulp, $MgSO_4$ 1 %

on pulp, treatment time 60 min, temperature 90°C, and O_2-pressure 5 bar. Set III consisted of 68 oxygen-delignified pulp samples from both conventional and displacement cooking of pine, Set IV 34 samples from both conventional and displacement cooking of birch, and Set V 20 oxygen-delignified displacement kraft birch pulp samples. The kappa number of the pulp samples (Table 4) was determined according to the SCAN-C 1:77 standard.

Table 4. Kappa numbers of the pulp samples[IV]

	SET I	**SET II**	**SET III**	**SET IV**	**SET V**
mean kappa no.	21,6	19,8	10,6	9,1	3,6
kappa no. min.	5,4	6,6	1,9	2,9	2,9
kappa no. max.	72	31,1	42,6	40	5,9
S*	11,3	6,4	8,8	7,9	1
%RSD**	52,4	32,3	83,4	87	27,8
total no. of samples	65	56	68	34	20

* S, standard deviation of kappa numbers within the set.
** %RSD, relative standard deviation of kappa numbers within the set (%).

Unbleached pulp samples with varying yields and kappa numbers (Table 5) were produced by laboratory-scale batch kraft cooking of Scots pine (*Pinus sylvestris*) and silver birch (*Betula verrucosa)*.[V] In addition, selected samples were further delignified using oxygen under the following conditions: consistency 5 %, NaOH charge 5 % on pulp, $MgSO_4$ 1 % on pulp, treatment time 60 min, temperature 90°C, and oxygen pressure 5 bar (Table 6).

Table 5. Kappa numbers and yields of unbleached pulp samples[V]

	KAPPA NUMBER		**YIELD**	
Pine	**calibration set**	**test set**	**calibration set**	**test set**
N*	53	46	53	46
min	7,8	8,3	33,9	35,7
max	27,4	27,4	44,8	44,7
mean	15,8	15,7	40,3	40,3
S**	5,6	5,6	2,9	2,8
Birch	**calibration set**	**test set**	**calibration set**	**test set**
N*	14	35	14	35
min	10,7	10,7	36,6	36,6
max	53,5	53,5	55,8	55,8
mean	20,7	18,5	46,6	46,3
S**	10,3	6,8	4,0	3,6

* N, number of samples.

** S, standard deviation.

Table 6. Kappa numbers and yields of oxygen-delignified pulp samples[V]

	KAPPA NUMBER		**YIELD**	
	calibration set	**test set**	**calibration set**	**test set**
N*	37	37	37	37
min	3,0	3,0	33,7	33,7
max	42,6	42,6	60,2	60,2
mean	13,1	13,1	43,6	43,6
S**	12,1	12,1	7,3	7,3

* N, number of samples.

** S, standard deviation.

4.2 PULP ANALYSIS

4.2.1 UV/Vis

A pulp sample was prepared for spectroscopic analysis according to the TAPPI T 218 om-91 standard.[I,IV-V] A diffuse reflectance spectrum ranging from 360 to 740 nm in even 10 nm intervals was measured using a properly calibrated Minolta CM 3700 dP spectrophotometer operating with the CIE standard illuminant D65/10°. The values from 400 to 740 nm were used in the data analysis. Each spectrum was the mean of three repeated measurements and at least two spectra corresponding to each sample were measured.

4.2.2 Other analysis

Digested pulp samples were thoroughly washed before screening.[I-V] The screened pulp samples were hydroextracted, homogenized, and their moisture content was determined with a Denver Instrument IR-200 moisture analyzer. The reject was oven-dried (105°C) to a constant mass. Oxygen-delignified pulp samples were washed, hydroextracted, and homogenized prior to a moisture-content determination.

4.3 ISOLATION OF THE LIGNIN

4.3.1 Black liquor

For the analytical determinations,[I-III] the kraft lignin sample was made as follows.[75] It was precipitated from the black liquor with 2 M hydrochloric acid (pH to 2), separated from the solution by centrifugation, washed thoroughly with acidified water, and finally air-dried at room temperature. For the Paper III, the preparation of black liquors and the determination of their combustion properties (*i.e.*, burning time and swelling) were performed elsewhere[76,77] and used as such.[III]

4.3.2 Sediments and receiving waters

Three sediment and six water samples from the kraft pulp mill watercourse near Äänekoski and Jyväskylä towns were collected at locations shown in Figure 4.[VI] A total of 13 water samples were collected from the mechanical pulp mill watercourse near Lohja and Tammisaari towns at locations shown in Figure 5.[VII]

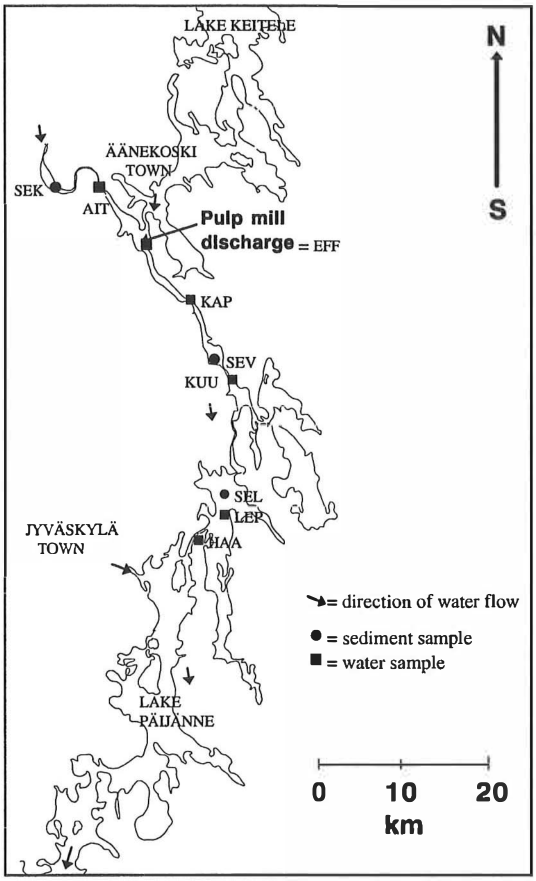

Figure 4. Map of the sampling area and the sampling sites.[VI]

The freeze-dried sediment sample (5 g) was extracted with 60 ml of 2 M NaOH solution and the extract was first neutralized with 2 M HCl and then filtered with a 0.45-µm membrane. The extract was ultrafiltrated using a Waters ultrafiltration device with two membranes having cutoff values of 1 and 10 kDa. After freeze-drying the fractions obtained with different membranes were weighed, combined, and homogenized. The water samples were adjusted to pH 10, filtered through a 0.45-µm membrane, and ultrafiltrated with 1 and 10 kDa membranes[VI] or with 1 kDa membrane only[VII] before freeze-drying, weighing, combining, and homogenizing.

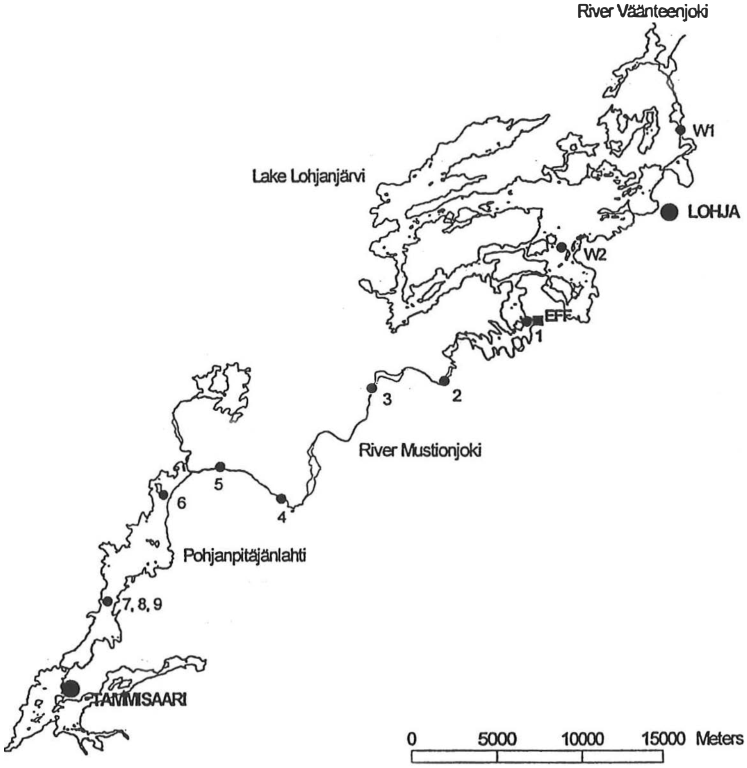

Figure 5. Map of the sampling area and the sampling sites.[VII]

4.4 ANALYSIS OF LIGNIN

4.4.1 NMR

Determination of the different hydroxyl groups in the kraft lignin samples was performed by ^{13}C and ^{31}P NMR spectroscopy.[II,III,78]

For the ^{31}P NMR measurements[II], kraft lignin samples were phosphitylated and their ^{31}P NMR spectra were measured using 2-chloro-1,3,2-dioxaphospholane as a phosphitylating reagent and 2,2-bis(4-hydroxyphenol)propane (bisphenol-A) as an internal standard. Measurements were carried out at 273 K with a Bruker Avance DRX 500 FT NMR spectrometer operating at 202.47 MHz using an inverse gated decoupling sequence.

^{13}C NMR spectra were taken from the acetylated kraft lignin samples (300 - 400 mg) in DMSO-d_6 on a JEOL GSX 270 spectrometer working at 67.9 MHz at 303 K with TMS as reference.[III] An average of 7000 - 16000 scans were collected with a 90° flip angle and a delay time of 9 s using an inverse gated decoupled pulse sequence. The selected integration ranges were: 169.8 - 171.8 ppm for primary acetoxy groups, 168.8 - 169.8 ppm for secondary acetoxy groups, and 166.0 - 168.8 ppm for phenolic acetoxy groups.

For the ^{31}P NMR measurements[III], the kraft lignin sample was phosphitylated according to Argyropoulos[79] using benzoic acid as an internal standard. ^{31}P NMR spectra were obtained at 308 K with inverse gated decoupling on a JEOL GSX 270 spectrometer working at 109.4 MHz. 2500 scans were collected with a 90° flip angle using a 5 s relaxation delay. The NMR data were processed by means of PERCH software[80,81] and integration ranges were set according to the earlier published procedures.[79]

4.4.2 Py-GC/MSD

A Pyroprobe 1000 pyrolysis unit (CDS Analytical, Inc.) equipped with a platinium filament was interfaced to GC/MSD system (Hewlett-Packard 5890 Series II GC with HP 5972 MSD).[VI] The pyrolysis temperature was 650°C for 10 s (heating-up rate 20°C/ms), interface temperature was 250°C, injector temperature 250°C, and detector temperature 280°C. Helium was used as a carrier gas at a rate of 50 ml/min for pyrolysis and 1.6 ml/min for the GC (split ratio was 1/30). Prior to pyrolysis, the sample was flushed with helium for 30 s. GC column was HP 5-MS (length 30 m, ID 0.25 mm, and film thickness 0.25 μm). The temperature program was 45°C (4 min) -->240°C (4°C/min) --> 280°C (39°C/min). Electron impact mass spectra were measured from 40 to 500 amu and ionization energy of 70 eV was used. Identification of the pyrolysis products was based on the mass spectra of the model compounds or data from the spectral library (Wiley 138) along with the retention times.

4.4.3 SEC

The molecular mass distribution and the average molecular masses (M_n, M_w, and M_z) of the kraft lignin samples (1 - 2 mg)[II-III] and the high-molecular-mass lignin material[VI] were analyzed using a HPSEC system consisting of a Waters 501 pump, a Waters 996 photodiode array detector, a Waters pump control module, a Waters 717 autosampler and a Waters column oven attached to a temperature control module. Four Waters Styragel HR columns (7.8 mm x 300 mm) were coupled in consecutive order according to pore size (HR 3 -> HR 2 -> HR 1 -> HR 0.5). Isocratic elution with tetrahydrofuran was carried out at a flow rate of 1.0 mL/min at 30°C. The UV spectra were measured in the range 225 - 500 nm at a rate of one spectrum/second with a resolution of 2.4 nm. In each case UV absorption at wavelength 280 nm was used for calculations of the average molecular masses. The molecular mass (MM) calibration was done with a set of polystyrene standards (MMs 280000 (V_0), 24150, 12860, 4075, and 2727 Da), along with bierol (MM 638 Da), dehydroisoeugenol (MM 326 Da), vanillin (MM 152 Da), and acetone (MM 58 Da).

4.5 ANALYSIS OF THE OXIDATIVE DEGRADATION PRODUCTS OF LIGNIN

4.5.1 HPLC

The cupric oxide oxidations[I,VI-VII] and the liquid chromatographic analysis of the oxidation products[VI-VII] were performed according to the method described elsewhere.[82]

4.5.2 GC/MSD

The gas chromatographic/mass spectroscopic analysis of the silylated sample solution was made with GC/MSD system (Hewlett-Packard 5890 Series II GC with HP 5972 MSD).[VII] The injector temperature was 260°C and detector temperature was 280°C. Helium was used as a carrier gas at a rate of 1.6 ml/min (split ratio was 1/30). GC column was Restek Rtx-5 MS (length 30 m, ID 0.25 mm, and film thickness 0.25 µm). The temperature program was 80°C (4 min) -->100°C (4°C/min) -->200°C (2°C/min) --> 280°C (15°C/min). The collected mass spectra range was 40 - 500 amu and ionization energy of 70 eV was used. Identification of the oxidation products was based on the mass spectra of the model compounds or data from the spectral library (Wiley 138) along with the retention times.

4.5.3 CZE

The CuO-oxidation products were analyzed with a Hewlett-Packard 3D electroproresis device using a fused silica capillary (L_{tot} 79.0 cm, L_{eff} 70.5 cm, 75 µm ID).[I] The 10 mmol borate buffer (pH 9.2) was prepared by mixing equimolar solutions of $Na_4B_2O_7$ and H_3BO_3, filtered through a 0.45-µm membrane and degassed in an ultrasonic bath prior to use as an electrolyte. The methanol solutions of the oxidation products were injected by a hydrodynamic injection (50 mbar, 6 s). A voltage of +18 kV was applied to the capillary (positive polarity, detector to cathode) during the 25 min run. The electropherograms were recorded at 210 nm. Before each analysis the capillary was first rinsed for 4 min with 1 M NaOH, then for 3 min with ultrapure water, and finally for 6 min with electrolyte solution. The

CuO-oxidation products identified were, in migration order, as follows: syringol, guaiacol, acetosyringone, acetovanillone, vanillin, 4-hydroxybenzaldehyde, and vanillic acid. Identification was based on the UV spectra of each component and the migration time of the corresponding model compound.

4.6 MULTIVARIATE DATA ANALYSIS

The capillary zone electrophoretic (CZE)[I] and chromatographic data from cupric oxide oxidation[VI,VII] and pyrolysis[VII], diffuse reflectance UV/Vis spectral[IV,V], ^{13}C and ^{31}P NMR spectral[III], along with size exclusion chromatographic and ^{31}P NMR spectral[II] data were each subjected (as the **X** matrix) to the PCA using a singular value decomposition (SVD). Prior to the multivariate calculations, the data were normed to the unit integral[II,VI], mean-centered[III-V], or autoscaled[I-III,VI-VII]. The regression models (PCR[IV,V] and PLS[I,III-V]) were calculated using either SVD or NIPALS for the mean-centered or autoscaled dependent variable or variables (the **Y** matrix). The rank of the models was determined by cross-validation[83] or by using separate validation set. The coefficient of determination (R^2) and standard error of prediction (SEP)[84] values were used to evaluate the models.

In the case of point-for-point 2D cross-correlation (2DCC) calculations[II], four PCA models (pine NMR, pine SEC, birch NMR, and birch SEC) were first computed (from the untreated data) and the data then reconstructed from the significant principal components (PCs), two in case of the NMR data and one in case of the SEC data. The coefficient of determination for the models in each case was 95.10% and 95.16% for the pine and birch NMR, respectively, and 98.83% and 99.68% for the pine and birch SEC, respectively.

A 2DCC calculation[18,85,86] was applied to these combined NMR (pine plus birch) and SEC (pine plus birch) PC datareconstructs and absolute coefficient values above 0.7 were taken into account. The correlation coefficients (r_{ij}) were calculated by the equation (1).[18]

Here, x_j is the SEC, and y_i the NMR PC datareconstructs. All computations were carried out on a personal computer using the MATLAB® software package.[87]

5 RESULTS AND DISCUSSION

5.1 UV/VIS DATA ON KRAFT PULPS

5.1.1 Determination of kappa number

A number of principal components was calculated from the mean-centered diffuse reflectance UV/Vis data on the 167 pulp samples.[IV] As can be seen from Figure 6, the principal component analysis suggested that separate models should be constructed for unbleached (U) and oxygen-treated (B) pulps.[IV] The variable loadings for the calculated principal components (Figure 7) also showed that wavelengths around 457 nm, used typically in the determination of pulp brightness (*e.g.*, ISO 2469, ISO 2470, TAPPI T 452 om-92, TAPPI T 525 om-92, and SCAN-P 3:93), were not the most differentiating ones. The first PC, explaining 96.5 % of variance, had the highest variable loading values clearly above 500 nm.[IV]

PCR and PLS models for the brownstock (Sets I and II) and bleached (Sets III, IV, and V) pulp samples were calculated and their predictive abilities were tested using separate test sets (Tables 7 and 8).[IV] Based on these data it could be concluded that the PLS models are slightly more accurate than the corresponding PCR models. In addition, the wood feedstock and the pulping method (conventional *vs.* displacement) had a strong effect on the kappa number prediction, as seen in Figures 8 and 9.

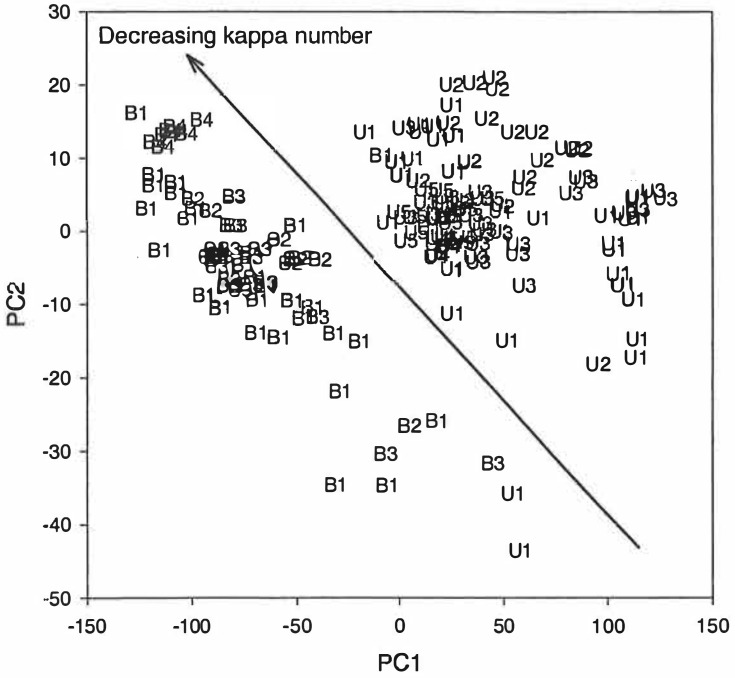

Figure 6. Score plot of the first and second Pcs.[IV]
U1 = conventional pine kraft; U2 = conventional birch kraft; U3 = displacement pine kraft; U4 = displacement birch kraft; U5 = conventional spruce kraft; B1= oxygen-delignified conventional pine kraft; B2 = oxygen-delignified conventional birch kraft; B3 = oxygen-delignified displacement pine kraft; B4 = oxygen-delignified displacement birch kraft.

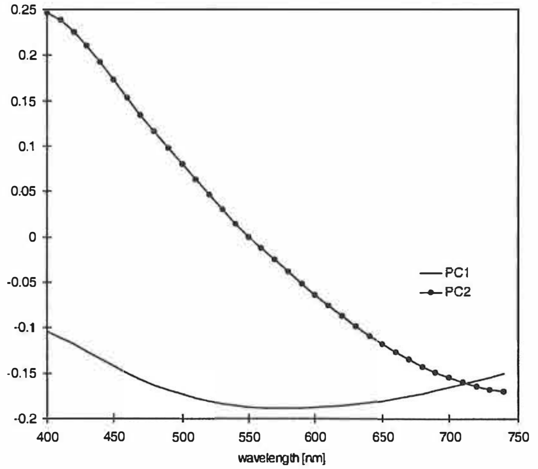

Figure 7. Variable loadings for the first and second PCs (in arbitrary units).[IV]

Table 7. Coefficient of determination (R^2, in cumulative %) for the test sets II and III[IV]

	Set II (brownstock pulp)		Set III (bleached pulp)	
Rank	PLS	PCR	PLS	PCR
1	3,5	2,1	74,0	73,6
2	50,7	48,7	79,2	78,9
3	56,7	54,9	81,0	80,9
4	85,6	63,1	84,0	81,1
5	87,5	62,6	88,5	81,1
6	90,3	86,7	89,0	88,1
7	89,8	87,8	91,8	88,4
8	90,8	89,0	92,8	88,5
9	74,7	93,5	94,5	91,2
10	81,0	93,9	94,6	92,7

Table 8. Coefficient of determination (R^2, in cumulative %) for the test sets IV and V[IV]

	Set IV (bleached pulp)		Set V (bleached pulp)	
Rank	PLS	PCR	PLS	PCR
1	73,5	73,4	55,5	50,9
2	75,4	74,8	76,6	75,6
3	77,3	76,5	80,4	77,4
4	83,7	79,5	89,7	87,8
5	93,6	87,4	99,5	98,6
6	95,9	92,8	99,9	99,2
7	97,8	95,0	ND*	ND
8	98,6	95,9	ND	ND
9	98,9	98,0	ND	ND
10	99,4	98,3	ND	ND

*ND, not determined.

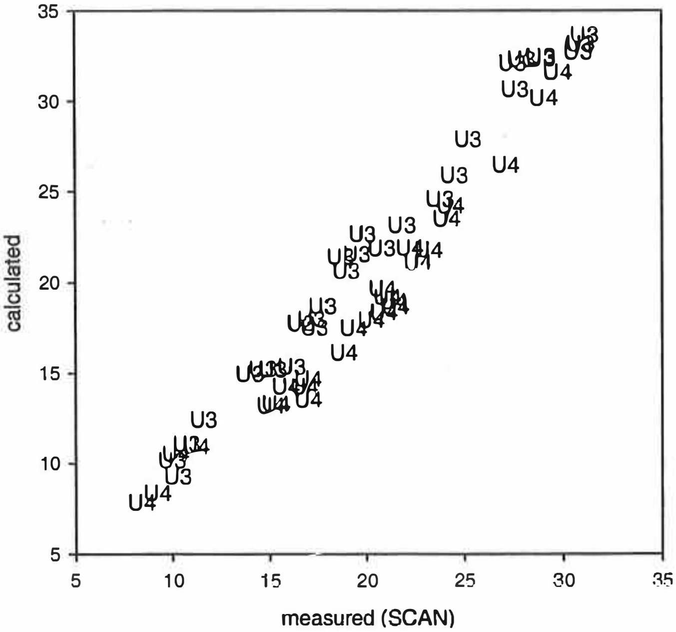

Figure 8. Measured kappa number *vs.* calculated kappa number for Set II.[IV] For abbreviations, see Figure 6.

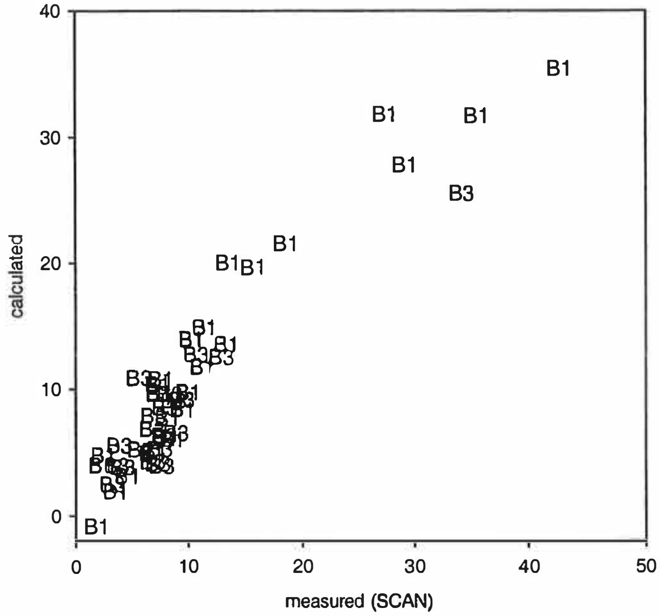

Figure 9. Measured kappa number *vs.* calculated kappa number for Set III.[IV] For abbreviations, see Figure 6.

5.1.2 Determination of pulping yield

To evaluate the validity of the diffuse reflectance UV/Vis data measured for this application the principal component analysis of the mean-centered spectral data of 256 pulp samples was performed.[V] The resulting PCA scores clearly differentiated the samples into separate categories depending on the delignification method. Figure 10 illustrates how the samples were classified according to the first and second principal coordinates, (the variance explained by each PC is presented in parenthesis). It can be concluded that separate models for kraft pulps with and without oxygen-alkali delignification are needed.

PCA also suggests that wood feedstock (softwood *vs.* hardwood) has a clear influence on modelling. Especially kraft pulps without oxygen delignification were sensitive for the feedstock and these pulps were grouped according to their wood feedstock origin.

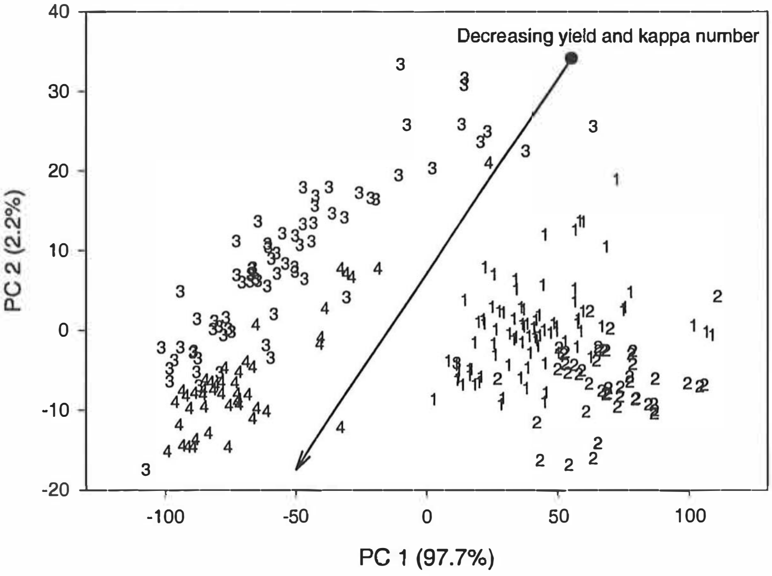

Figure 10. Score plot of the first and second PCs.[V]
1 = pine kraft; 2= birch kraft; 3 = oxygen-delignified pine kraft;
4 = oxygen-delignified birch kraft.

Separate PLS models were constructed for pine and birch kraft pulps. The R^2 and SEP values of the test set are presented in Table 9. The model for the pine pulps predicted kappa number very well (R^2 more than 92 % and SEP less than 2 units), whereas the yield prediction was slightly less accurate, giving the R^2 and SEP values around 90 % and less than one unit, respectively. The model for the birch samples gave a very good prediction of yield and a fairly good prediction of kappa number (the R^2 values 93 % and 80 %, respectively).

One PLS model was shown to be sufficient for oxygen delignified pulps.[V] The model obtained was thus based on both the pine and birch samples. The R^2 and SEP values of the test set are presented in Table 10. The kappa number prediction functioned extremely well, giving R^2 values above 98 % and SEP values around one unit. The yield prediction was also good, resulting high R^2 values (above 94 %) and low SEP values (less than two units).

Table 9. The rank, coefficient of determination (R^2, in cumulative %), and standard error of prediction (SEP, in original units) for the test set of the PLS model for the kraft pulp samples without oxygen delignification[V]

	PINE				BIRCH			
	Kappa number		Yield		Kappa number		Yield	
Rank	R^2	SEP	R^2	SEP	R^2	SEP	R^2	SEP
1	34,2	20,6	28,6	5,8	(-53,1)	(68,1)	13,7	11,6
2	90,5	3,0	88,5	1,1	67,3	14,6	69,4	4,3
3	89,7	3,2	87,2	1,2	66,0	15,1	88,0	1,9
4	90,3	3,0	89,4	0,9	70,8	13,0	92,5	1,3
5	92,9	2,2	90,0	0,9	80,0	8,9	93,6	1,2
6	94,2	1,8	90,0	0,9	80,6	8,7	93,9	1,1
7	94,5	1,7	90,9	0,8	69,9	13,4	90,2	1,6
8	95,0	1,6	93,2	0,6	76,9	10,3	90,3	1,6
9	95,6	1,4	93,2	0,6	74,7	11,3	85,9	2,2
10	95,6	1,4	93,7	0,6	75,2	11,1	61,9	5,3

Table 10. The rank, coefficient of determination (R^2, in cumulative %), and standard error of prediction (SEP, in original units) for the test set of the PLS model for the kraft pulp samples with oxygen delignification[V]

	Kappa number		Yield	
Rank	R^2	SEP	R^2	SEP
1	96,6	4,8	80,0	10,4
2	97,0	4,2	80,1	10,4
3	98,1	2,6	81,7	9,6
4	98,2	2,6	85,7	7,5
5	99,2	1,2	94,7	2,8
6	99,2	1,1	95,6	2,3
7	99,3	0,9	95,8	2,2
8	99,3	0,9	96,9	1,6
9	99,3	0,9	97,5	1,3
10	99,3	0,9	98,5	0,8

The variable loading values of all the calculated models indicated that wavelengths around 457 nm, used typically in the determination of pulp brightness (*e.g.*, ISO 2469, ISO 2470, TAPPI T 525 om-92, and SCAN-P 3:93), are not the most differentiating ones. The loadings of the first three latent predictors are presented as a function of the wavelength for unbleached and oxygen-delignified pulps in Figures 11 and 12, respectively.

Wavelengths from 400 nm to 450 nm were important for the models, especially in case of unbleached pulps, but the most informative area was around 600 nm. Due to the nature of this type of models, there was, however, no need to reduce the amount of data prior to PLS modelling. Therefore, the use of a wide wavelength range is strongly recommended.

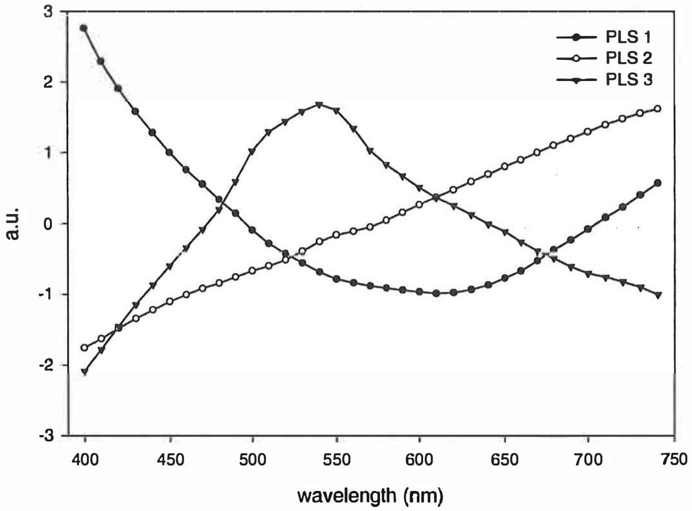

Figure 11. PLS model for the unbleached kraft pulps: the variable loadings for the first, second, and third latent variables.[V]

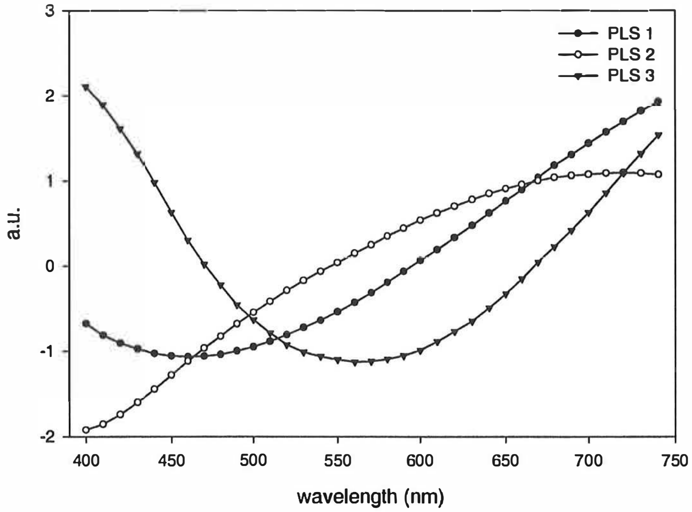

Figure 12. PLS model for the oxygen-delignified kraft pulps: the variable loadings for the first, second, and third latent variables.[V]

5.2 KRAFT LIGNIN

5.2.1 Structural changes in lignin during pulping

^{31}P NMR and SEC data were subjected to the PCA using a singular value decomposition. Prior to the analysis, the NMR data (20 spectra, 1400 points each) were mean-centered and scaled to the unit variance. The SEC data (20 chromatograms, measured at 245 nm, 770 points each) were normed to the unit integral.[II]

The score plots, describing the classification of the samples in the principal coordinates (explained variance of each principal component is presented in parenthesis), are presented in Figures 13 - 16.[II]

The results obtained from both types of data were in close agreement and yielded a valid classification of the samples, thereby confirming the applicability of this kind of analysis. In addition, both sets of results indicated feedstock-related differences in the delignification rate, seen in the different groupings of the pine and birch samples.[II]

A 2DCC image (GPC data as abscissa and ^{31}P NMR data as ordinate) with a correlation threshold of 0.7 is presented in Figure 17.[II] Although it is safe to conclude that changes in the lignin molecular mass within the mass range 1000 - 3200 Da are highly correlated to changes in all types of lignin structural units, it is also clear that the only marked correlations in the higher-molecular-mass region were with resonances around 135 and 137 ppm. These resonances are typical of those of *threo*- and *erythro*-β-aryl ether structures, respectively.[86]

Cross-correlations around 133 ppm were due to the impurities in the phosphitylating reagent. There were also two thin cross-correlation lines at 131.4 and 131.6 ppm that could be attributed to a minor molecular mass increase induced by the condensed lignin structures (approximative chemical shift range 131.2 - 131.9 ppm).[II]

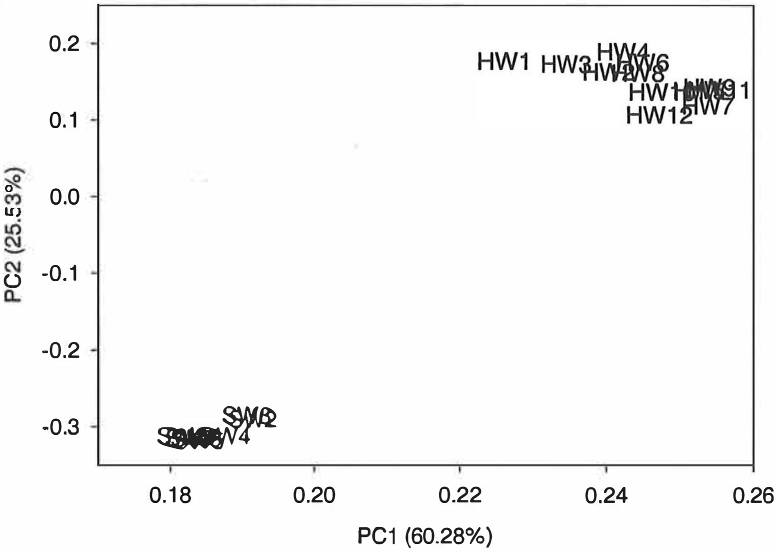

Figure 13. Score plot of the first and second PCs, NMR data.[II] For abbreviations, see Table 2.

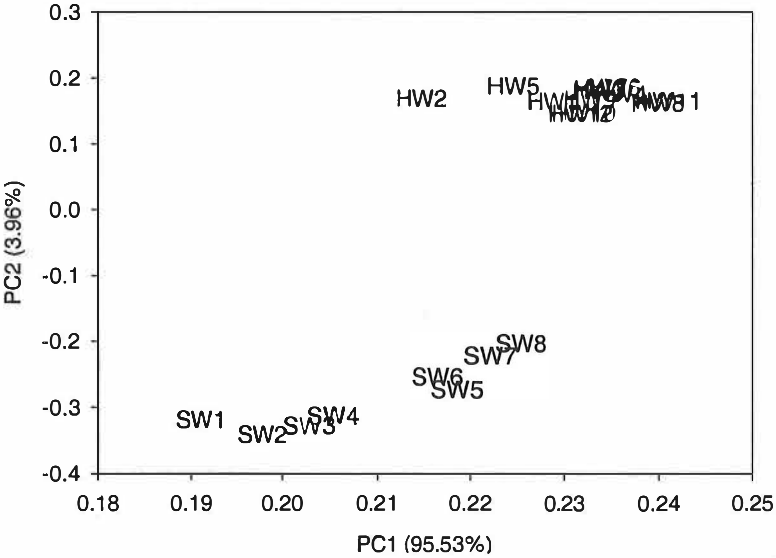

Figure 14. Score plot of the first and second PCs, SEC data.[II] For abbreviations, see Table 2.

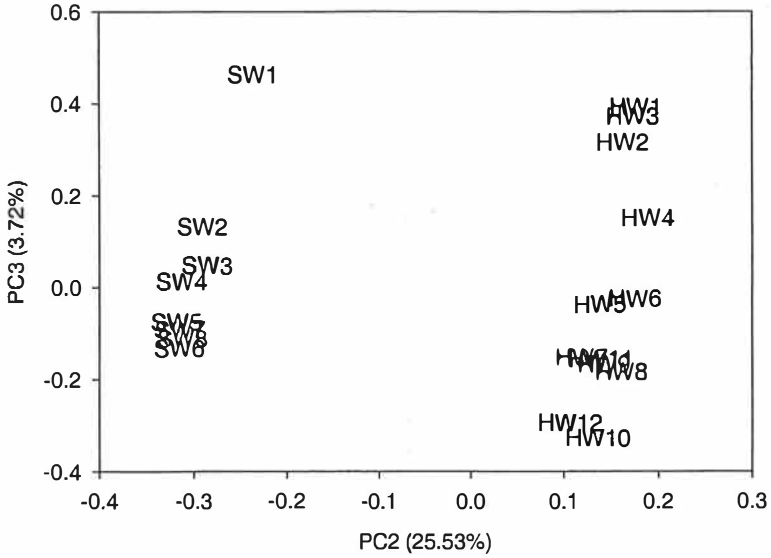

Figure 15. Score plot of the second and third PCs, NMR data.[II] For abbreviations, see Table 2.

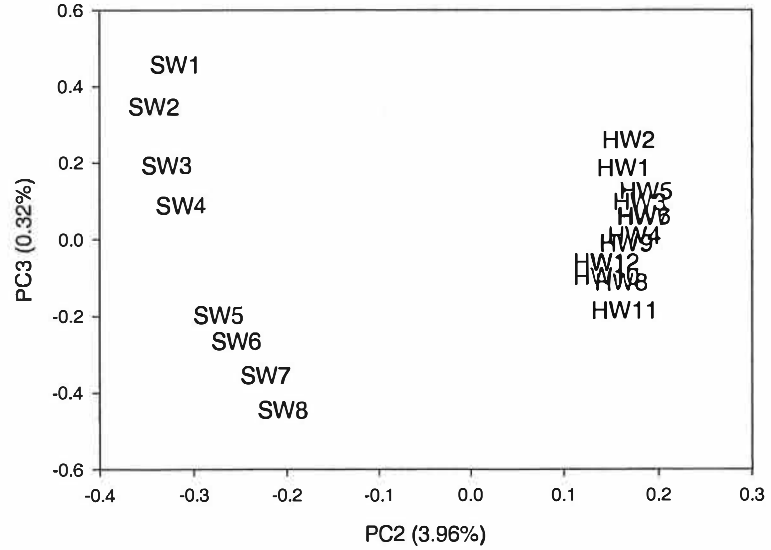

Figure 16. Score plot of the second and third PCs, SEC data.[II] For abbreviations, see Table 2.

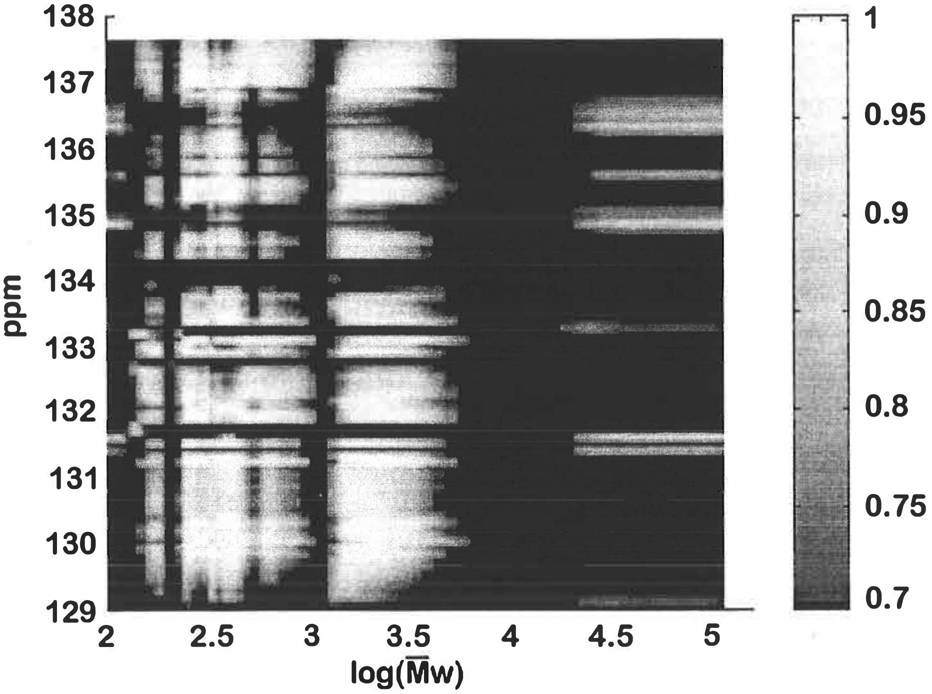

Figure 17. 2D cross-correlation image with a correlation threshold of 0.7.[II]

5.2.2 Indirect modelling of kraft pulp properties

For the PCA models, a number of principal components was calculated using a singular value decomposition of the autoscaled CZE data (peak heights).[I] The PCR and PLS models were calculated from the autoscaled CZE data (peak heights) of the identified and four most prominent unidentified kraft lignin CuO-oxidation products as the **X** matrix and the autoscaled total cooking yield, kappa number, and ISO brightness values (from the corresponding pulp samples) as the **Y** matrix. A NIPALS algorithm (PLS2) was used in these calculations.

A biplot of the PCA model, describing both the classification of the samples and the variable loadings on the principal coordinates, is presented in Figure 18.[I] The model indicated that samples SW1 and SW7 might differ significantly from the others. Also, the loading values suggested the existence of a synergistic effect on the part of oxidation products sharing a similar type of aromatic ring structure. Therefore, another PCA model with 9 new interaction variables (based on this structural similarity) along with the original variables was constructed using a NIPALS algorithm. The interaction variables were created by multiplying certain variable pairs from the autoscaled CZE data (peak heights) and adding these to the original **X** matrix. These pairs were: syringol-acetosyringone, guaiacol-acetovanillone, guaiacol-vanillin, guaiacol-vanillic acid, acetovanillone-vanillin, acetovanillone-vanillic acid, vanillin-vanillic acid, acetovanillone-acetosyringone, and acetosyringone-vanillic acid. This modified model (interaction variables included) clearly classified the kraft lignin samples according to the delignification method, with the exception of sample SW1. Also, samples SW2 and SW7 showed slightly different score values for the first two PCs, SW7 being markedly different in the third PC. The score plot of the model with the interaction variables is presented in Figure 19. The R^2 values for the PCA models are shown in Table 11.[I]

Table 11. The rank and coefficient of determination (R^2, in cumulative %) for the PCA models[I]

Rank	PCA model	PCA model with interactions
1	65,9	53,9
2	84,1	76,6
3	91,2	86,1
4	95,9	94,2
5	99,0	97,3

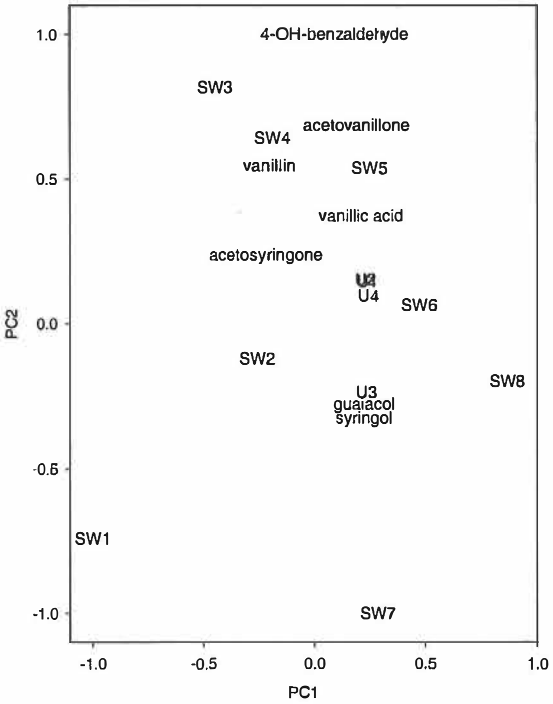

Figure 18. Biplot of the PCA model.[I] U1 - U4 refer to the four most prominent unidentified compounds. For other abbreviations, see Table 1.

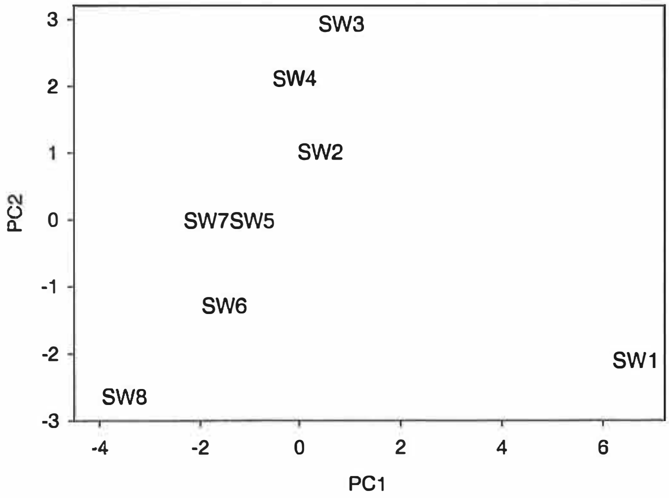

Figure 19. Score plot of the PCA model with interaction variables.[I] For abbreviations, see Table 1.

Both PCR and PLS models for the total cooking yield, kappa number, and ISO brightness values of the kraft pulp were constructed. The outlying samples SW1 and SW7, sharing approximately similar experimental errors to other samples, were chosen to be kept in the models mainly to give as wide a range of Y variable space as possible. The R^2 and SEP values for both the PCR and PLS models in significant rank are shown in Table 12. Another set of PCR and PLS models were calculated using the interaction variables along with the original variables. In these models the sample SW7 was not considered as an outlier, at least according to its Hotelling T^2 value. The sample SW1 was also retained in the models to give as wide a range to the Y variables as possible. The sample SW1 had, however, approximately similar leverage and residual X variance in these models than the other samples, except the sample SW2, which had the smallest levarage and largest X residual variance. The R^2 and SEP values for these models in significant rank are shown in Table 13.[I]

On the basis of these data it could be concluded that the use of interaction variables in constructing both types of models clearly increased their modelling capabilities, as seen in the increase in R^2 and decrease in SEP. The oxidative cleavage of lignin polymers leads mainly to the oxidation products with a similar aromatic backbone, either of the guaiacyl or syringyl type. These products in turn seemed to intercorrelate in a multivariate manner not expressable in univariate or bivariate models. Several combinations of interaction variables were tested for the purpose and the results confirmed that the choice of variables should be made very carefully on the basis of structural similarity, with the execption of the variable formed using values for acetosyringone and vanillic acid. For this application, the best combination of variables (at least in terms of cross-validated coefficient of determination and SEP values) was the one reported here.[I]

The main difference between the PCR and the PLS models was in the number of latent variables needed to create a significant model. Whilst the PCR models constructed for these data needed 4 principal components, the corresponding PLS models were calculated using 3 latent variables in the case of models including the interaction variables. The measured *vs.* calculated values for the PLS models (with interaction variables) for the total cooking yield, kappa number, and ISO brightness are presented in Figures 20 - 22.[I]

Table 12. The coefficient of determination (R^2, in cumulative %) and standard error of prediction (SEP, in original units) values of the PCR and PLS models[I]

Model	PCR			PLS		
	R^2	SEP	rank	R^2	SEP	rank
Total cooking yield	98,1	0,941	4	93,8	1,698	2
Kappa number	96,9	3,840	4	98,6	2,530	4
ISO brightness	97,3	0,584	4	95,1	0,779	2

Table 13. The coefficient of determination (R^2, in cumulative %) and standard error of prediction (SEP, in original units) values of the PCR and PLS models with interaction variables[I]

Model	PCR			PLS		
	R^2	SEP	rank	R^2	SEP	rank
Total cooking yield	99,1	0,660	4	99,4	0,514	3
Kappa number	99,9	0,689	4	99,9	0,385	3
ISO brightness	98,3	0,462	4	98,9	0,367	3

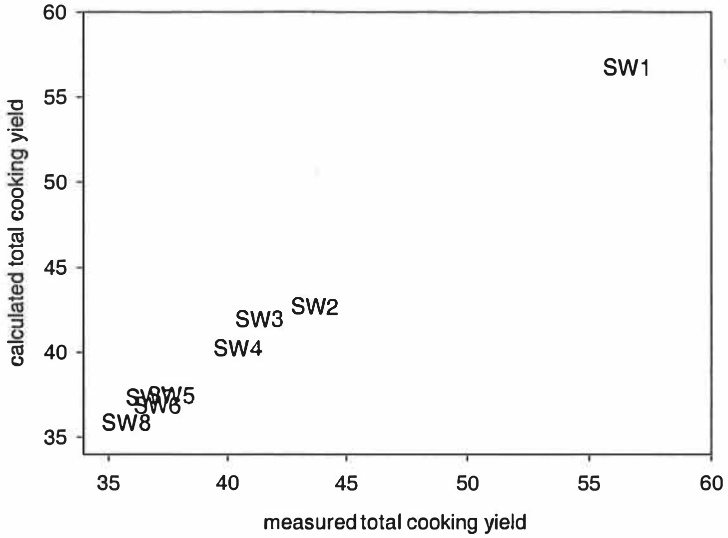

Figure 20. Measured *vs.* calculated total cooking yield for softwood pulps.[I] For abbreviations, see Table 1.

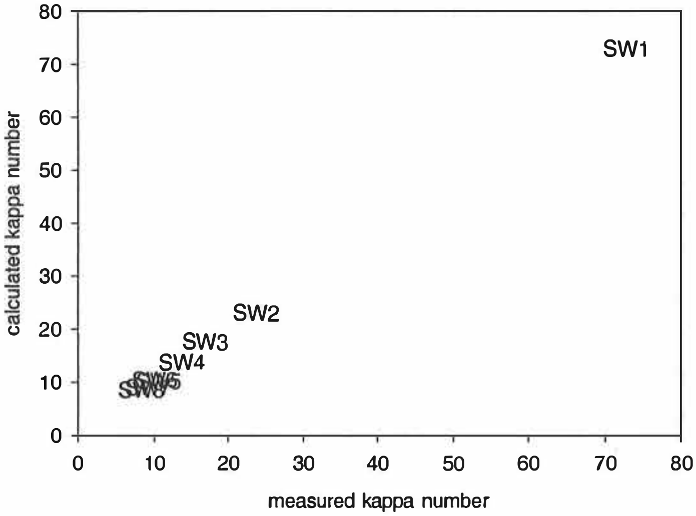

Figure 21. Measured *vs.* calculated kappa number for softwood pulps.[I] For abbreviations, see Table 1.

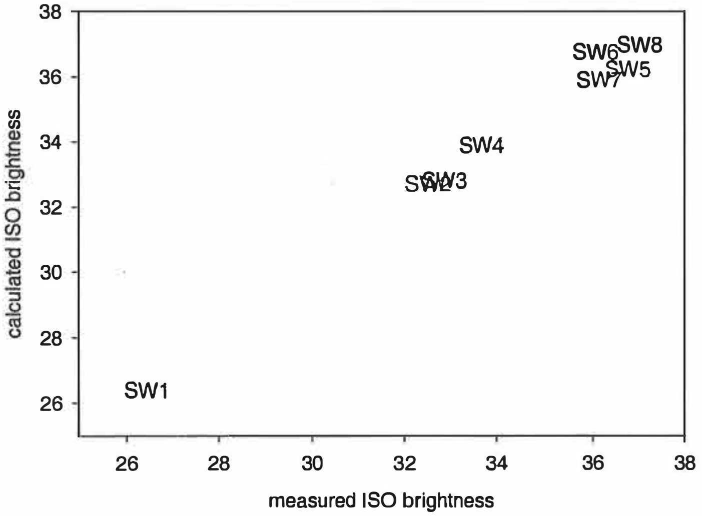

Figure 22. Measured *vs.* calculated ISO brightness for softwood pulps.[I] For abbreviations, see Table 1.

5.3 CORRELATION BETWEEN LIGNIN STRUCTURE AND THE COMBUSTION PROPERTIES OF BLACK LIQUOR

Multivariate calculations were performed for the following data sets.[III] Set 1 ("analysis data") consisted of the NMR and SEC results (total of 12 independent variables) along with combustion properties as dependent variables. Set 2 ("spectral data") consisted of the ^{31}P NMR spectral data (20 spectra, 813 points each) as independent variables and the combustion properties as dependent variables. Typical spectra of phosphitylated pine and birch kraft lignins are presented in Figure 23.

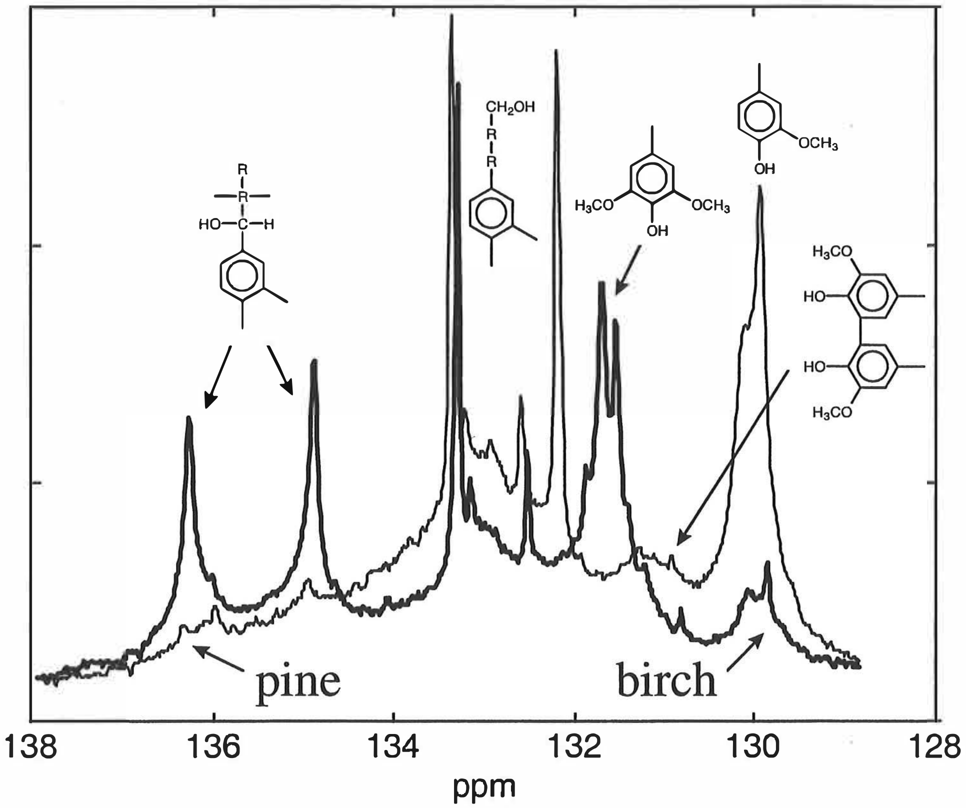

Figure 23. Typical ^{31}P NMR spectra of the phosphitylated pine and birch kraft lignins.[III]

For Set 1, the PCA model consisting of two principal components clearly assigned the black liquor samples into separate groups with respect to feedstock material (PC 1) and cooking time (PC 2). A score plot of the PCA model is presented in Figure 24. The PLS models predicted both the burning time and swelling of black liquor with reasonable accuracy (Tables 14 and 15).

For Set 2, the PCA model with two principal components explained over 92 % of the total variance and assigned the black liquor samples according to the feedstock material. The variable loading values plotted against the original ppm-scale could be interpreted as sub-spectra and they yielded detailed information on changes in the original spectra of the samples which were otherwise very difficult to observe.[III]

The variable loadings for the first PC represented the mean spectrum of the samples, whereas the variable loadings for the second and third principal components, illustrated in Figure 25, indicated the presence of carbohydrate together with lignin fragments. Sharp peaks between 132 and 134 ppm, which could be assigned to the carbohydrate units, were directly related to the less resolved peaks originating from the phenolic hydroxyls (129.1 - 132.0 ppm) in the lignin structure.[III]

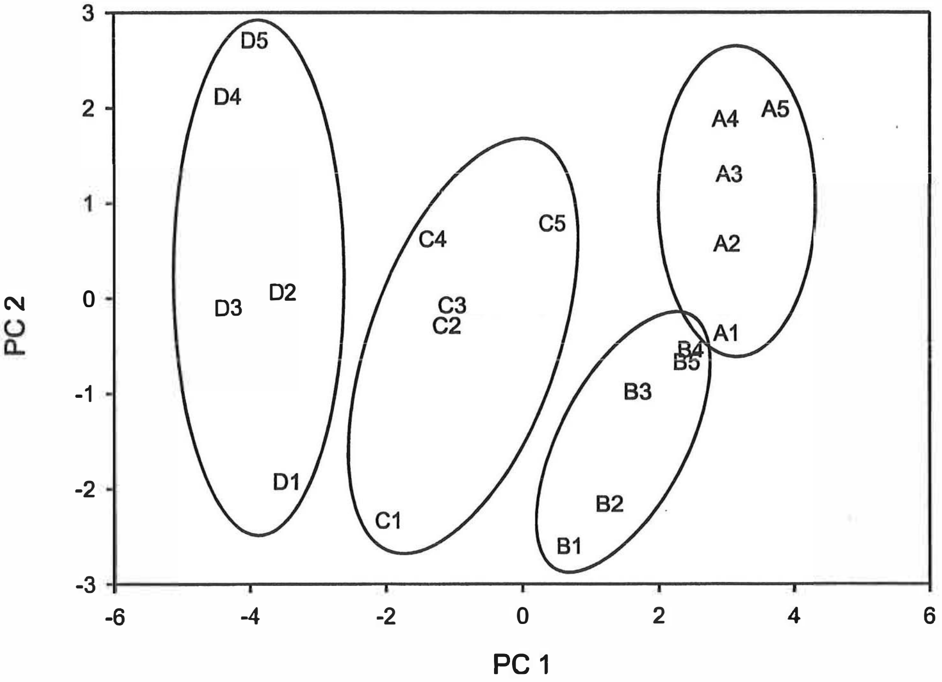

Figure 24. Score plot of the PCA model for Set 1. For abbreviations, see table 3.[III]

Moreover, at least two different types of guaiacyl hydroxyls present in "lignin-carbohydrate complexes" were distinguished on the basis of the difference in chemical shift in the guaiacyl-OH region (129.1 - 131.5 ppm) between the second and third PC. The PLS model for burning time performed with reasonable accuracy (Tables 14 and 15), showing relatively low SEP values. Swelling of black liquor droplets was accurately predicted with the PLS model.[III]

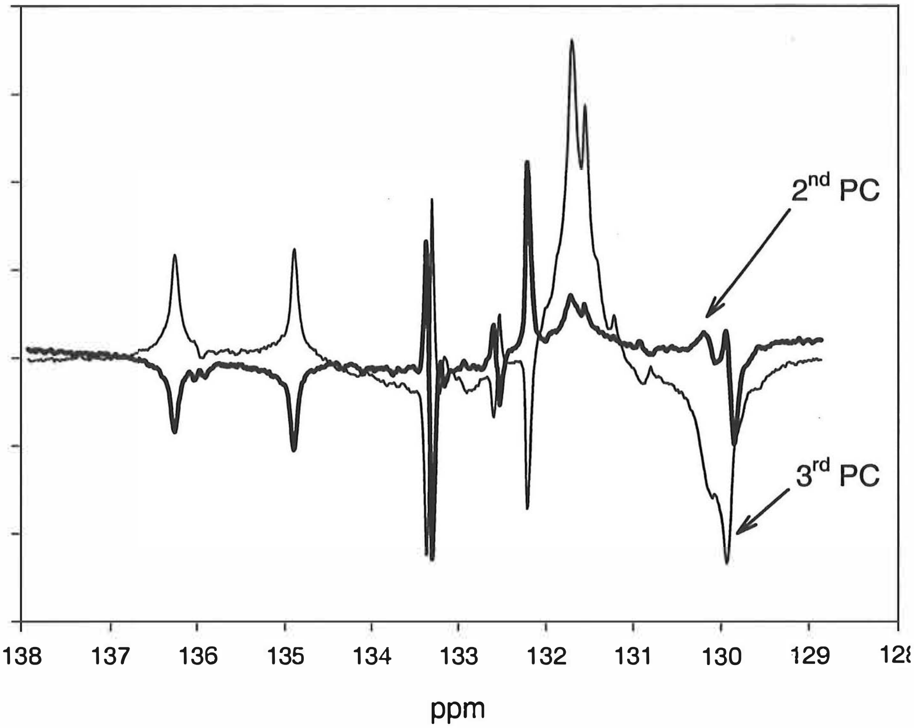

Figure 25. The variable loading values for the second and third PCs.[III]

Table 14. The rank, coefficient of determination (R^2), and standard error of prediction (SEP, in original units) for the PLS models for burning time[III]

	Set 1		Set 2	
Rank	R^2	SEP	R^2	SEP
1	62,6	0,0555	43,1	0,0045
2	74,8	0,0456	46,6	0,0042
3	79,8	0,0408	57,8	0,0033
4	81,3	0,0393	77,6	0,0018
5	82,2	0,0384	82,0	0,0014
6	82,8	0,0376	90,0	0,0008
7	83,1	0,0373	90,2	0,0008
8	83,6	0,0368	90,9	0,0007

Table 15. The rank, coefficient of determination (R^2), and standard error of prediction (SEP, in original units) for the PLS models for swelling[III]

	Set 1		Set 2	
Rank	R^2	SEP	R^2	SEP
1	33,2	4,79	36,4	20,75
2	54,4	3,96	51,7	15,78
3	71,8	3,12	57,7	13,82
4	76,8	2,82	75,1	8,13
5	77,1	2,80	84,6	5,02
6	78,0	2,75	96,9	1,02
7	78,5	2,72	98,4	0,51
8	78,7	2,70	99,4	0,19

5.4 CHARACTERIZATION OF LIGNIN IN RECEIVING WATERS

5.4.1 Chemical pulp mill

Structures of the identified compounds formed during CuO-oxidation of lignin material[VI VII] are presented in Figure 26.

The results from multivariate calculations of CuO-oxidation data[VI] are presented in Figure 27 (the variance explained by each PC is presented in parenthesis). In the picture the scores (representing samples) and loadings (representing the effect of each CuO-oxidation product) of the second and third PCs are presented. As expected, it can be seen that the chlorinated oxidation products (originating from the chlorine dioxide bleaching) were differentiating the effluent sample (EFF) from the others. In addition, the background sample (AIT) was clearly different from the others, mostly due to the 4-hydroxyphenyl moieties. One sample (HAA) contained large amount of phthalic acid (PHTAC) which could readily be seen from the groupings in Figure 27.

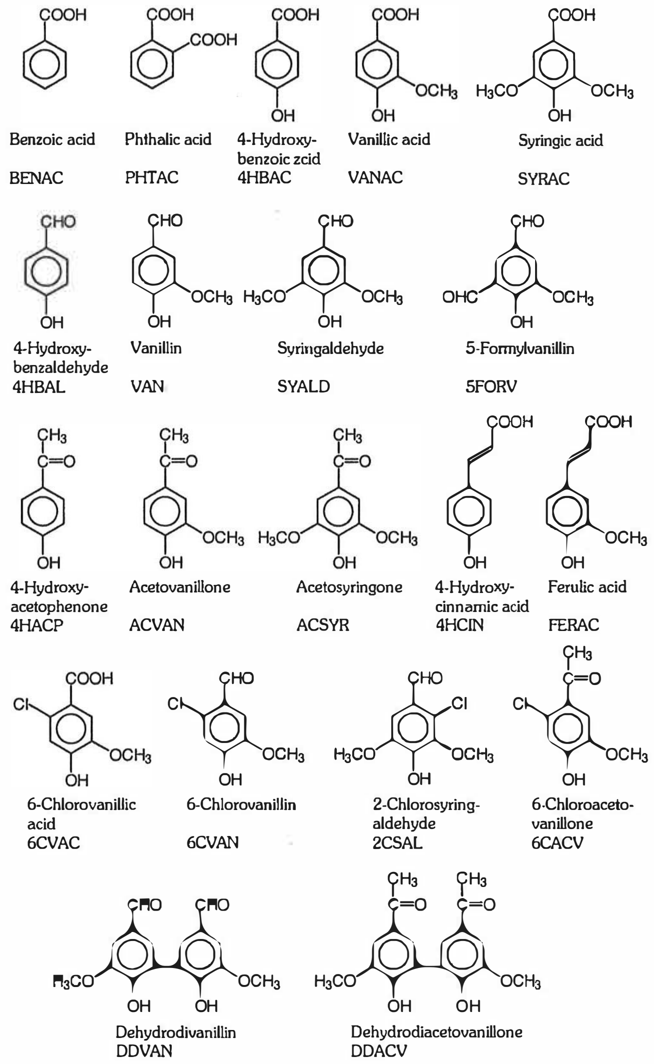

Figure 26. Structures of the identified CuO-oxidation products.[VI,VII]

The PCA calculated from the pyrolysis data clearly grouped the sediment and water samples of which the effluent and background samples differed markedly.[VI] The scores values of the second and third principal components are presented in Figure 28 (the variance explained by each PC is presented in parenthesis).

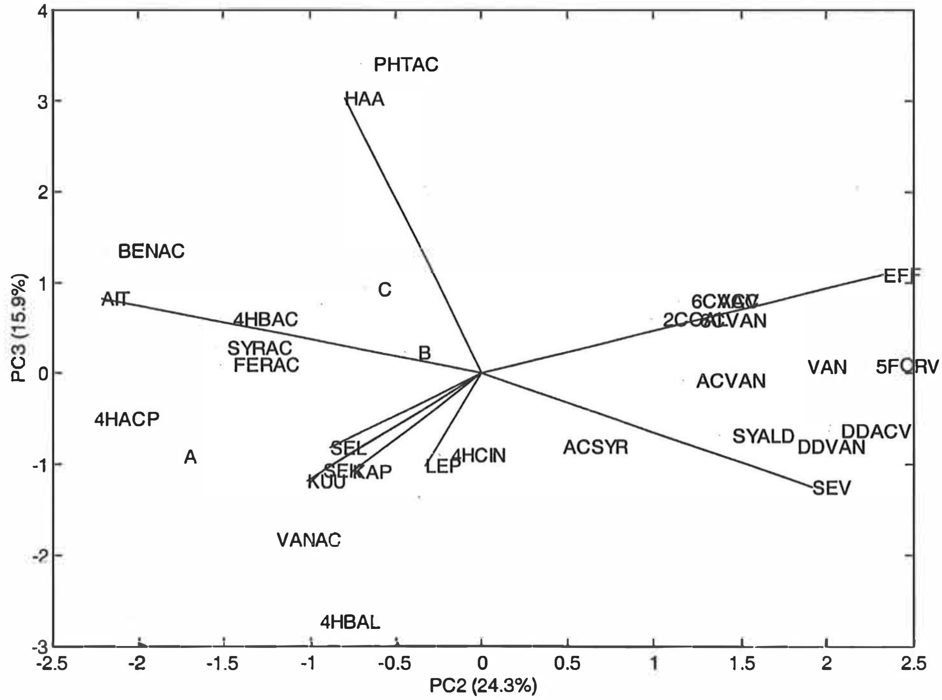

Figure 27. Biplot of the second and third PCs, CuO-oxidation data.[VI] A, B, and C are unknown oxidation products. For other abbreviations, see Figures 4 and 26.

5.4.2 Mechanical pulp mill

The PCA model calculated from CuO-oxidation data[VII] unambiguously classified the background samples (W1 - 2), effluent samples (EFF1 - 4), river water samples (1 - 5), and sea water samples (7 - 10) into separate groups.

Sample 6 was an intermediate sample which was not directly grouped to any class. The background samples differed from the others mainly due to the higher amount of the 4-hydroxyphenyl and syringyl moieties. A biplot of the PCA model is presented in Figure 29 (the variance explained by each PC is presented in parenthesis).[VII]

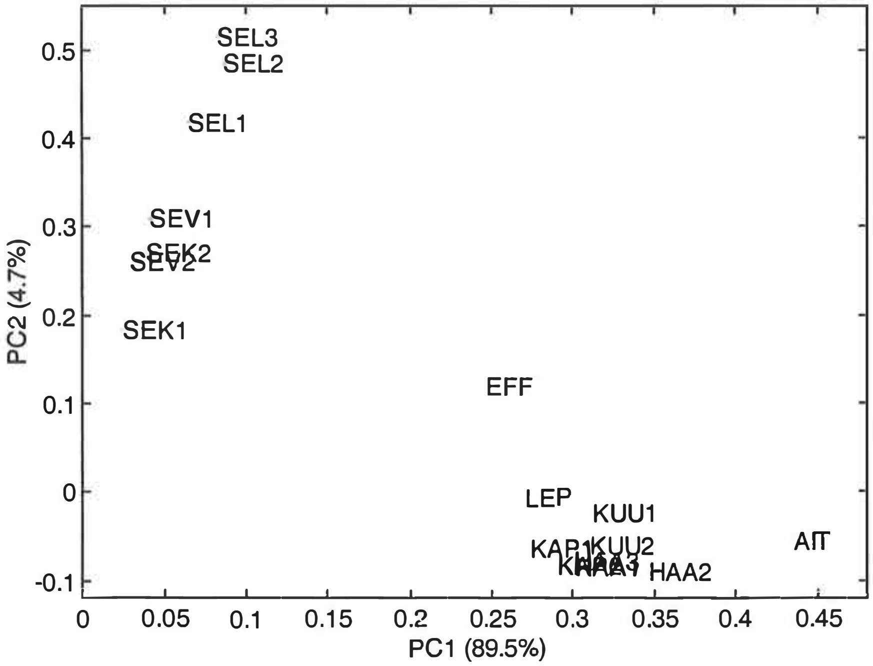

Figure 28. Score plot of the first and second PCs, pyrolysis data.[VI] For abbreviations, see Figures 4 and 26.

In addition, vanillin (VAN), dehydrodivanillin (DDVAN), and dehydrodiacetovanillin (DDACV) had also a rather large influence on the PCA model, seen as high variable loading values especially in the second PC (Figure 29) .[VII]

The number of PCs (rank) and the corresponding coefficient of determination values (R^2) for the PCA models calculated from the data on high-molecular-mass lignin material are presented in Table 16.[VI,VII]

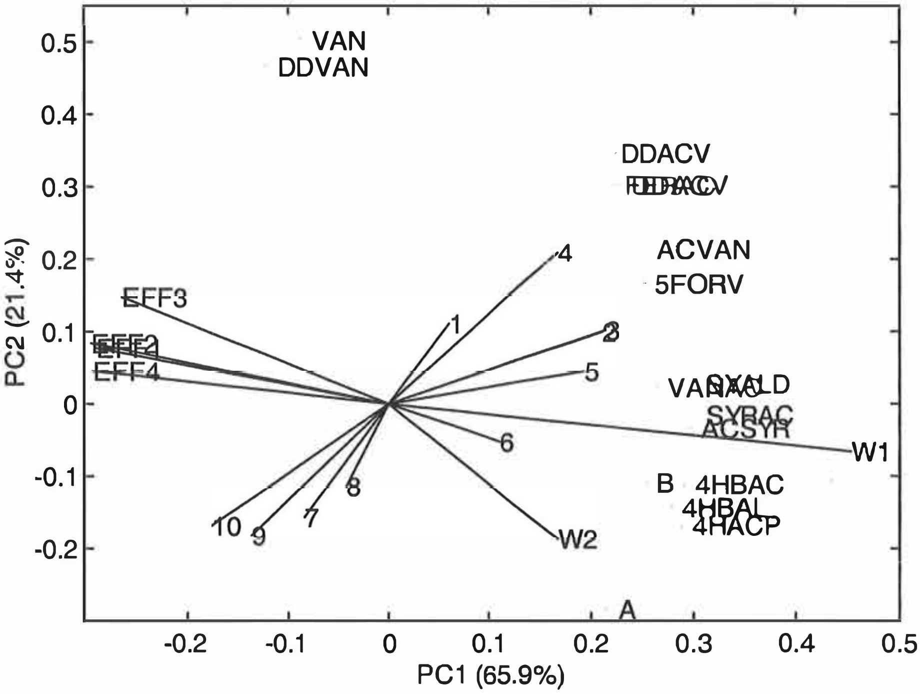

Figure 29. Biplot of the first and second PCs, CuO-oxidation data.[VII] A, B, and C are unknown oxidation products. For other abbreviations, see Figures 5 and 26.

Table 16. The number of principal components (rank) and coefficient of determination (R^2, in cumulative %) for the PCA models[VI,VII]

Rank	CuO-Oxidation-PCA[VI]	Pyrolysis-PCA[VI]	CuO-Oxidation-PCA[VII]
1	31,8	89,5	65,9
2	56,1	94,2	87,3
3	72,0	97,4	92,0
4	83,0	98,4	95,3
5	90,8	98,9	97,3
6	96,2	99,2	98,7
7	99,0	99,4	99,2

6 CONCLUSIONS

The chemometric approach based on several relevant analytical procedures and measurements provided a versatile tool for encountering research challenges in the wood processing chemistry. The interpretation of inherently complex and ambiguous empirical datasets obtained in this work was greatly facilitated using multivariate data analysis.

The principal component analysis, for example, made possible the classification of different types of kraft lignins and pulps, including detailed investigation of their spectral features. Effective noise reduction of chromatographic and spectroscopic data was performed using principal component calculations.

Multivariate regression modelling (PCR and PLS) of ISO brightness, kraft pulp yield, kappa number, and the combustion properties of kraft black liquors was accomplished. These types of models could be regarded as a sound basis for the more advanced models and applications (so-called “chemometric sensors”) in practical use. In addition, two-dimensional cross-correlation calculations could be applied as a source of detailed information, for example, to the case where simultaneous changes in lignin structure and molecular mass were detected.

7 REFERENCES

1 Anon. *FAO Factbook, Preliminary 1998 Forestry data. http://apps.fao.org/cgi-bin/nph-db.pl?subset=forestry* [30.11.1999].

2 Sjöström, E. *Wood Chemistry - Fundamentals and Applications*, 2nd edn., Academic Press, San Diego, USA, 1993.

3 Minor, J. In: Dence, C. and Reeve D. (Eds.) *Pulp Bleaching - Principles and Practise*, TAPPI Press, Atlanta, USA, 1996, pp. 27-57.

4 Rydholm, S. Pulping Process, John Wiley & Sons, London, UK, 1965, pp. 335-337.

5 Frederick, Wm. In: Adams, T. (Ed.) *Kraft Recovery Boilers*, TAPPI Press, Atlanta, USA, 1997, pp. 61-99.

6 Grace, T. In: Grace, T. and Malcolm, E. (Eds.) *Pulp and Paper Manufacture Vol. 5 - Alkaline Pulping,* 3rd edn., TAPPI Press, Atlanta, USA, 1989, pp. 473-478.

7 Fletcher, D. In: Grace, T. and Malcolm, E. (Eds.) *Pulp and Paper Manufacture Vol. 5 - Alkaline Pulping,* 3rd edn., TAPPI Press, Atlanta, USA, 1989, pp. 609-615.

8 Clayton, D., Easty, D., Einspahr, D., Lonsky, W., Malcolm, E., McDonough, T., Schroeder, L., and Thompson, N. In: Grace, T. and Malcolm, E. (Eds.) *Pulp and Paper Manufacture Vol. 5 - Alkaline Pulping,* 3rd edn., TAPPI Press, Atlanta, USA, 1989, pp. 114-128.

9 Hase, A., Koppinen, S., Riistama, K., and Vuori, M. *Suomen Kemianteollisuus*, 4th edn., Chemas Oy, Tampere, Finland, 1998, pp. 126-138.

10 McDonough, T. In: Dence, C. and Reeve D. (Eds.) *Pulp Bleaching - Principles and Practise*, TAPPI Press, Atlanta, USA, 1996, pp. 215-239.

11 Bokström, M. and Nordén, S. *Proc. 1998 Int. Pulp Bleaching Conf.*, Helsinki, Finland, pp. 23-32.

12 Reeve, D. In: Dence, C. and Reeve D. (Eds.) *Pulp Bleaching - Principles and Practise*, TAPPI Press, Atlanta, USA, 1996, pp. 3-24.

13 Lorås, V. In: Casey, J. (Ed.) *Pulp and Paper - Chemistry and Chemical Technology*, 3rd edn., vol 1, John Wiley & Sons, New York, USA, 1980, p. 663.

14 Martens, H. and Næs, T. *Multivariate Calibration*, John Wiley & Sons, Chichester, UK, 1989.

15 Jolliffe, I. *Principal Component Analysis*, Springer, New York, USA, 1986.

16 de Jong, S. *Chemom. Intell. Lab. Syst.* **18** (1993) 251.

17 Bradley, D. *Scientific Computing World* **44** (1998) 29.

18 Barton, F., Himmelsbach, D., Duckworth, J., and Smith, M. *Appl. Spectrosc.* **46** (1992) 420.

19 Wold, S., Antti, H., Lindgren, F., and Öhman, J. *Chemom. Intell. Lab. Syst.* **44** (1998) 175.

20 Booksh, K. and Kowalski, B. *Anal. Chem.* (1994) **66** 782A.

21 Faber, N., Buydens, L., and Kateman, G. *J. Chemom.* **8** (1994) 147.

22 Faber, N., Buydens, L., and Kateman, G. *J. Chemom.* **8** (1994) 181.

23 Faber, K., Lorber A., and Kowalski, B. *J. Chemom.* **11** (1997) 95.

24 Booksh, K., Lin, Z., Wang, Z., and Kowalski, B. *Anal. Chem.* (1994) **66** 2561.

25 Bro R. *Chemom. Intell. Lab. Syst.* **42** (1997) 149.

26 Dahl, K., Piovoso, M., and Kosanovich, K. *Chemom. Intell. Lab. Syst.* **46** (1999) 161.

27 Andersson, C. and Bro, R. *Chemom. Intell. Lab. Syst.* **45** (1998) 93.

28 Bro, R. *J. Chemom.* **10** (1996) 47.

29 Antalek, B., Hornak, J., and Windig W. *J. Magn. Reson.* **132** (1998) 307.

30 Wallbäcks, L., Edlund, U., Nordén, B., and Iversen, T. *Nord. Pulp Pap. Res. J.* **6** (1991) 74.

31 Michell, A. *Appita* **47** (1994) 29.

32 Wallbäcks, L., Edlund, U., Lindgren, T., and Agnemo, R. *Nord. Pulp Pap. Res. J.* **10** (1995) 88.

33 Schimleck, L., Michell, A., and Vinden, P. *Appita* **49** (1996) 319.

34 Lennholm, H., Rosenqvist, M., Ek, M., and Iversen, T. *Nord. Pulp Pap. Res. J.* **9** (1994) 10.

35 Malkavaara, P., Isoaho, J.-P., Alén, R., and Soininen, J. *J. Chemom.* accepted.

36 Michell, A. and Schimleck, L. *Appita* **49** (1996) 23.

37 Lennholm, H., Wallbäcks, L., and Iversen, T. *Nord. Pulp Pap. Res. J.* **10** (1995) 46.

38 Michell, A. and Schimleck, L. *Appita* **51** (1998) 127.

39 Seisto, A., Poppius-Levlin, K., and Fuhrmann, A. *Pap. Puu* **80** (1998) 520.

40 Broderick, G., Paris, J., Valade, J., and Wood, J. *Pap. Puu* **77** (1995) 410.

41 Lindgren, T., Edlund, U., and Iversen, T. *Cellulose* **2** (1995), 273.

42 Lennholm, H., Larsson, T., and Iversen, T. *Carbohydr. Res.* **261** (1994) 119.

43 Wallbäcks, L., Edlund, U., Nordén, B., and Berglund, I. *Tappi J.* **74** (1991) 201.

44 Schimleck, L., Wright, P., Michell, A., and Wallis, A. *Appita* **50** (1997) 40.

45 Lennholm, H. and Iversen, T. *Holzforschung* **49** (1995) 462.

46 Olsson, R., Tomani, P., Karlsson, M., Josefsson, T., Sjöberg, K., and Björklund, C. *Tappi J.* **78** (1995) 158.

47 Rodrigues, J., Faix, O., and Pereira, H. *Holzforschung* **52** (1998) 46.

48 Antti, H., Sjöström, M., and Wallbäcks, L. *J. Chemom.* **10** (1999) 591.

49 Meder, R., Gallagher, S., Mackie, K., Böhler, H., and Meglen, R. *Holzforschung* **53** (1999) 261.

50 Thygesen, L. *Holzforschung* **50** (1996) 434.

51 Thygesen, L. *J. Near Infrared Spectrosc.* **2** (1994) 127.

52 Borga, P., Hämäläinen, M., and Theander, O. *Holzforschung* **46** (1992) 299.

53 Brunner, M., Eugster, R., Trenka, E., and Bergamin-Strotz, L. *Holzforschung* **50** (1996) 130.

54 Hoffmeyer, P. and Pedersen, J. *Holz als Roh- und Werkstoff* **53** (1995) 165.

55 Furumoto, H., Lampe, U., Meixner, H., and Roth, C. *Papier*, **53** (1999) 235.

56 Marklund, A., Hauksson, J., Edlund, U. and Sjöström, M. *Nord. Pulp Pap. Res. J.* **14** (1999) 140.

57 Marklund, A., Hauksson, J., Edlund, U. and Sjöström, M. *Nord. Pulp Pap. Res. J.* **13** (1998) 211.

58 Vanchinathan, S., Ananth, S., Jarvis, J., and Krishnagopalan, G. *Tappi J.* **79** (1996) 187.

59 Lingren, T. and Edlund U. *Nord. Pulp Pap. Res. J.* **13** (1998) 76.

60 Schimleck, L. and Michell, A. *Tappi J.* **81** (1998) 229.

61 Benar, P., Goncalves, A., Mandelli, D., Ferreira, M., Schuchardt, U. *J. Wood Chem. Technol.* **19** (1999) 151.

62 Evaggeli, B., Koutsoula, E., and Koukios, E. *Proc. 5th Eur. Workshop on Lignocellul. Pulp*, University of Aveiro, Aveiro, Portugal, 1998, 475.

63 Evaggeli, B. Argyropoulos, D., and Koukios, E. *Proc. 5th Eur. Workshop on Lignocellul. Pulp*, University of Aveiro, Aveiro, Portugal, 1998, 183.

64 Danielsson, L.-G. and Sheng C. *Process Control Qual.* **6** (1994) 149.

65 Peralta-Zamora, P., Dezotti, M., Reyes, J., and Duran, N. *Proc. 4th Braz. Symp. Chem. Lignins Other Wood Components*. Universidade Federal de Pernambuco, Recife, Brazil, 1995, 77.

66 Ledung, L., Ulmgren, P., and Hupa, M. *Nord. Pulp Pap. Res. J.* **12** (1997) 145.

67 Heikka, R. and Minkkinen, P. *Pap. Puu* **75** (1993) 592.

68 Heikka, R. and Minkkinen, P. *Pap. Puu* **76** (1994) 154.

69 Teppola, P., Mujunen, S.-P. and Minkkinen, P. *Chemom. Intell. Lab. Syst.* **38** (1997) 197.

70 Mujunen, S.-P., Teppola, P., Minkkinen, P., and R.-S. Wirkkala, *Chemom. Intell. Lab. Syst.* **41** (1998) 83.

71 Teppola, P., Mujunen, S.-P., and Minkkinen, P. *Chemom. Intell. Lab. Syst.* **41** (1998) 95.

72 Teppola, P., Mujunen, S.-P., and Minkkinen, P. *Chemom. Intell. Lab. Syst.* **45** (1999) 371.

73 Teppola, P. Mujunen, S.-P., Minkkinen, P., Puijola, T., and Pursiheimo, P. *Chemom. Intell. Lab. Syst.* **44** (1998) 313.

74 Teppola, P., Mujunen, S.-P., and Minkkinen, P. *Chemom. Intell. Lab. Syst.* **45** (1999) 23.

75 Alén, R. and Hartus, T. *Cellul. Chem. Technol.* **22** (1988) 613.

76 Whitty, K., Backman, M., Forssén, M., Hupa, M., and Sorvari, V. *J. Pulp Paper Sci.* **23** (1997) J117.

77 Alén, R., Louhelainen, J., Malkavaara, P., and Rantanen, K. *Proc. 1998 TAPPI International Chemical Recovery Conference*, Tampa, USA, Vol. Three, 1105.

78 Malkavaara, P. and Alén, R. *Proc. 8th International Symposium on Wood and Pulping Chemistry*, Helsinki, Finland, (1995), Vol. III, 103.

79 Argyropoulos, D. *J. Wood Chem. Technol.* **14** (1994) 45.

80 Laatikainen, R., Niemitz, M., Weber, U., Sundelin, J.-P., Hassinen, T., and Vepsäläinen, J. *J. Magn. Reson.* **A120** (1996) 1.

81 Laatikainen, R., Niemitz, M., Malaisse, W., Biesemans, M., and Willem, R. *Magn. Reson. Med.* **36** (1996) 359.

82 Hyötyläinen, J., Knuutinen, J., and Vilén, E. *Chemosphere* **30** (1995) 891.

83 Wold, S. *Technometrics* **20** (1978) 397.

84 Geladi, P., MacDougall, D., and Martens, H. *Appl. Spectrosc.* **39** (1985) 491.

85 Malkavaara, P., Toivanen, T.-J., and Alén, R. *J. Chemom.* accepted.

86 Malkavaara, P., Alén, R., and Kolehmainen, E. *Proc. 10th International Symposium on Wood and Pulping Chemistry*, Yokohama, Japan, (1999), Vol. II, 158.

87 MATLAB® for Windows®, version 4.2c, The MathWorks, Inc., Natick, USA, 1994.

Paper I

Petteri Malkavaara, Pasi Harjula, Raimo Alén, and Juha Knuutinen, Chemometric Investigation on Structural Changes in Pine Kraft Lignin During Pulping, *Chemom. Intell. Lab. Syst.* accepted.

https://doi.org/10.1016/S0169-7439(00)00090-3

Paper II

Petteri Malkavaara, Raimo Alén, and Erkki Kolehmainen, Chemometrics: An Important Tool for the Modern Chemist; An Example from Wood Processing Chemistry, *J. Chem. Inf. Comp. Sci.* **40(2)** (2000) 438-441.

https://doi.org/10.1021/ci990444i

Paper III

Petteri Malkavaara, Raimo Alén, and Erkki Kolehmainen, Multivariate Correlation between ^{13}C and ^{31}P NMR Spectral Data on Dissolved Lignin and the Combustion Properties of Kraft Black Liquor, *Magn. Reson. Chem.* **37** (1999) 407-412.

https://doi.org/10.1002/(SICI)1097-458X(199906)37:6<407::AID-MRC475>3.0.CO;2-S

Paper IV

Petteri Malkavaara and Raimo Alén, A Spectroscopic Method for Determining Lignin Content of Softwood and Hardwood Kraft Pulps, *Chemom. Intell. Lab. Syst.* **44(1-2)** (1998) 287-292.

https://doi.org/10.1016/S0169-7439(98)00068-9

Paper V

Petteri Malkavaara and Raimo Alén, Diffuse Reflectance Vis and Chemometrics as a Useful Tool for Determining Pulp Yield and Kappa Number, *Tappi J.* accepted.

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2969368

Paper VI

Juha Hyötyläinen, Juha Knuutinen, Petteri Malkavaara, and Jussi Siltala, Pyrolysis-GC-MS and CuO-Oxidation-HPLC in the Characterization of HMMs from Sediments and Surface Waters Downstream of a Pulp Mill, *Chemosphere* **36(2)** (1998) 297-314.

https://doi.org/10.1016/S0045-6535(97)10006-6

Paper VII

Juha Hyötyläinen, Juha Knuutinen, and Petteri Malkavaara, Transport of High Molecular Mass Lignin Material in the Receiving Water System of a Mechanical Pulp Mill, *Chemosphere* **36(3)** (1998) 577-587.

https://doi.org/10.1016/S0045-6535(97)00366-4