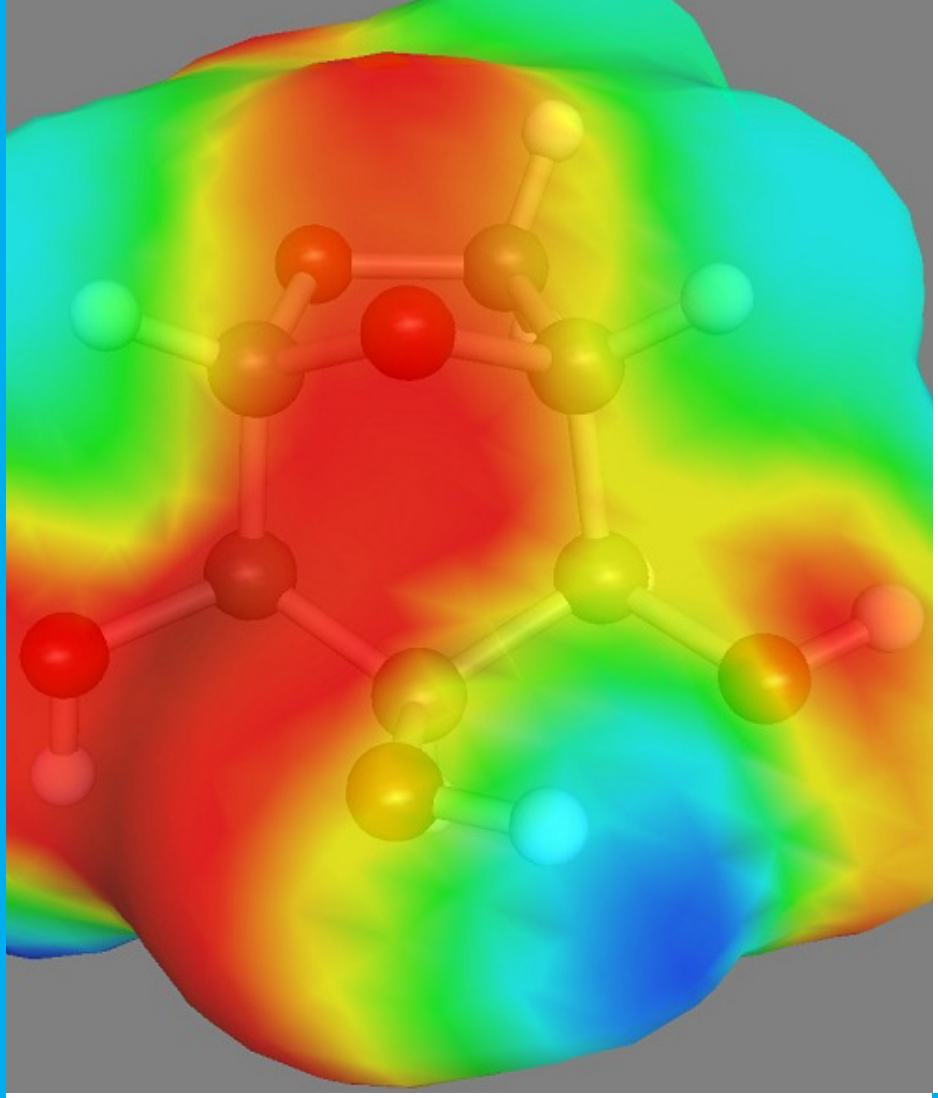

TECHNOLOGY • RESEARCH HIGHLIGHTS • VISIONS • SCIENCE

67

BIOSCEN

Modelling Biorefinery Scenarios

VTT TECHNOLOGY 67

BIOSCEN

Modelling Biorefinery Scenarios

Juha Leppävuori & Pertti Koukkari (eds.)

ISBN 978-951-38-7902-0 (URL: http://www.vtt.fi/publications/index.jsp)
ISSN 2242-122X (URL: http://www.vtt.fi/publications/index.jsp)

JULKAISIJA – UTGIVARE – PUBLISHER

VTT
PL 1000 (Tekniikantie 4A, Espoo)
02044 VTT
Puh. 020 722 111, faksi 020 722 4374

VTT
PB 1000 (Teknikvägen 4A, Esbo)
FI-02044 VTT
Tfn +358 20 722 111, telefax +358 20 722 4374

VTT Technical Research Centre of Finland
P.O. Box 1000 (Tekniikantie 4A, Espoo)
FI-02044 VTT, Finland
Tel. +358 20 722 111, fax + 358 20 722 4374

BIOSCEN
Modelling Biorefinery Scenarios

Biojalostamoprosessien mallinnus. **Juha Leppävuori & Pertti Koukkari (eds.)**
Espoo 2012. VTT Technology 67. 159 p. + app. 17 p.

Abstract

Biorefining unit processes were studied in terms of their physicochemical modelling properties and parameters, complemented by surrogate metamodelling and extensive sensitivity analysis. The unit processes included flash condensation of the fast biomass pyrolysis, thermal and enzyme catalysed hydrolysis of lignocellulosic biomass and fermentation of the hydrolysis product to bioethanol and biobutanol. In addition, a flowsheet based mass and energy balance was developed for the bark biorefinery, the key factors of which were then assessed by using the elementary efficiency method for sensitivity analysis. With the results from hydrolysis and fermentation models, a comparison of the greenhouse gas emissions from barley straw based ethanol and butanol was performed. The bark-based biorefinery producing mainly tannin and ethanol proved to be economically challenging. The results also show that biobutanol is in general more demanding (in terms of greenhouse gas emissions) than the production of bioethanol.

Keywords biorefinery, material properties, modelling, parameter optimisation, pyrolysis oil condensation, lignocellulosic biomass, hydrolysis, fermentation, greenhouse gas emissions

Biojalostamoprosessien mallinnus

BIOSCEN. Modelling Biorefinery Scenarios. **Juha Leppävuori & Pertti Koukkari (toim.).** Espoo 2012. VTT Technology 67. 159 s. + liitt. 17 s.

Tiivistelmä

Joukkoa biojalostamon yksikköprosesseja tutkittiin simulointimallien avulla. Tutkimuksessa tarkasteltiin tyypillisten biojalostamotuotteiden fysikaalis-kemiallisten ainearvojen saatavuutta ja näiden soveltamista termodynaamisissa ja kemialliseen reaktoritekniikkaan perustuvissa laskentamalleissa. Parametrien luotettavuutta testattiin metamallinnuksen ja herkkyysanalyysien avulla. Tutkitut yksikköprosessit olivat pyrolyysiöljyn talteenotto kondensoimalla, terminen ja entsyymikatalysoitu lignoselluloosan hydrolyysi sekä edelleen hydrolyysituotteiden fermentointi bioetanoliksi ja biobutanoliksi. Olkeen perustuvalle bioetanolin ja biobutanolin tuotannolle tehtiin mallinnukseen perustuva kasvihuonekaasupäästöjen analyysi. Tanniinia ja bioetanolia tuottavalle kuoribiojalostamolle laadittiin aine- ja energiataseet kuvaava prosessimalli, jonka taloudellisuutta tutkittiin edelleen parametrisen herkkyysanalyysin avulla.

Avainsanat biorefinery, material properties, modelling, parameter optimisation, pyrolysis oil condensation, lignocellulosic biomass, hydrolysis, fermentation, greenhouse gas emissions

Preface

The purpose of the BioScen project was to develope quantitative modelling approaches for the future biomass based processes producing fuels and chemicals. The aim of the BioScen project was in developing methods to estimate the necessary material properties for phase and reaction equilibria, for the calculation of unit processes and their integration to biorefining production plant simulations. Additional focus was laid on model optimisation and product life cycle analysis. As the biorefining technologies possess an extensive range from thermal pyrolysis to biochemical processing at ambient temperatures, a most generic thermodynamically based approach was selected to enable usage of the methods to the wide variety of possible applications. The methods were then applied to a number of case studies including modelling of flash condensation of pyrolysis oil, hydrolysis of cellulosic biomass and its product recovery and the subsequent fermentation processes for bioethanol and biobutanol, for which also a comparative life cycle analysis was performed. Flowsheet process simulation was applied to a conceptual wood bark biorefinery. Metamodelling techniques were used for both model and parameter optimisation, including their sensitivity analysis.

Modelling of the Biorefinery Scenarios, BIOSCEN, was a Tekes funded project under the BioRefine program. The research partners were Aalto University, The University of Jyväskylä and VTT Technical Research centre of Finland. The goal of the project was to develop an intellectual foundation for the modelling of biorefinery processes ranging from physical and chemical properties of key substances processes to flowsheet process optimization and life cycle analysis of selected biorefinery concepts. Thanks to active steering group, the project execution was intensively supported and evaluated.

The text and data contributors were Timo Aittokoski (particularly Chapters 2.2 and 2.3), Anna Zaytseva (Chapters 2.1, 3–4), Petteri Kangas (Chapter 5), Peter Blomberg (Chapters 6–8), Sakari Kaijaluoto and Jouni Savolainen (Chapter 9,) Reetta Sorsa (Chapters 2.4 and 10) and Juha-Pekka Pitkänen (Chapter 11), The script has been revised and edited by Juha Leppävuori and Pertti Koukkari who also have been contributing to Chapters 8 (Leppävuori) and 1, 2 and 11 (Koukkari).

Contents

Appendices

List of symbols

A_p	surface area of catalyst (particle)
c_A	molar concentration of substance A
D, D_A	diffusion coefficient (of substance A)
EOS	equation of state
Hi	Henry's law coefficient
k	reaction rate coefficient
K-value	distribution coefficient of substance between two separate phases
L_A	heat of vaporisation of component A
LLE	liquid-liquid equilibrium
M_A	molar mass
P, P_i	pressure and partial pressure of component *i*
P_i	parachor of component *i*
$\boldsymbol{r, r_A}$	rate of reaction
r_p	radius of a catalyst particle
T	temperature
T_r, P_r	reduced temperature and reduced pressure
V, V_i	volume and molar volume of component *i*
V_p	volume of a catalyst (particle)
VLE	vapour-liquid equilibrium
VLLE	vapour-liquid-liquid equilibrium
x_i	mole fraction of component *i*
BSA	bovine serum albumin
CBD	cellulose binding domain
CBH1	cellobiohydrolase 1, also Cel7A
CD	catalytic domain
Cel6A	non-progressive exoglucanase
Cel7A	progressive exoglucanase
Cel7B	endoglucanase
EB	enzyme (β-glucosidase) complex containing a cellobiose molecule
EG	enzyme (β-glucosidase) complex containing a glucose molecule
EG1	endoglucanase 1, also Cel7B

EGG	enzyme (β-glucosidase) complex containing two glucose molecules
SSF	simultaneous saccharification and fermentation
γ, γ_A	surface tension, of component A
λ	catalys distance-radius ratio
η, η_A	viscosity, of component A
ρ_A	bulk density of component A
ν_i	stoichiometric coefficient
$\Delta_f G$	Gibbs energy of formation
$\Delta_f G'$	Transformed Gibbs energy of formation
ΔH_{vap}	enthalpy of vaporisation
a,b,c	general EOS moel parameters
τ_{ij}	binary interaction parameter (COSMO)
Δu_{ij}	interaction energy difference (COSMO)
ω_i	compound specific Mathias-Copeman parameter (SRK-EOS)
a^{index}	reduced Helmholtz energy of index-specified properties (SAFT-EOS)
ε_{ij}	binary interaction energy (SAFT-EOS)
κ^{AiBi}	effective association volume (parameter of SAFT-EOS)
σ_{ij}	binary interaction parameter (SAFT-EOS)
E	total emissions from the use of the fuel
e_{ec}	emissions from the extraction or cultivation of raw materials
e_l	annualised emissions from carbon stock changes caused by land use change
e_p	emissions from processing
e_{td}	emissions from transport and distribution
e_u	emissions from the fuel in use
e_{sca}	emission saving from soil carbon accumulation via improved agricultural management
e_{ccs}	emission saving from carbon capture and geological storage
e_{ccr}	emission saving from carbon capture and replacement
e_{ee}	emission saving from excess electricity from cogeneration

1. Introduction

1.1 Project organisation

BioScen project was commenced in November 2008 as part of the TEKES funded BioRefine technology programme. The programme's expected total volume will reach about €250 million by 2012. A majority of the funding is focussed on enterprise projects, while the research project volume during the five-year period has totalled €41 million, of which Tekes' funding was about €25 million. The respective figures for the BioScen modelling project were €1.2 million and €0.84 million. The industry share in the funding of the BioScen project was 9%.

BioScen was organised as a joint project of Aalto University (Aalto), University of Jyväskylä (JYU) and VTT. The responsible leaders were professor Ville Alopaeus (Aalto), professor Kaisa Miettinen (JYU) and research professor Pertti Koukkari (VTT), who also acted as the coordinator of the project. Project manager was Dr. Juha-Pekka Pitkänen from VTT.

The three work packages of the project were

- physical, chemical and transport properties of biorefinery chemicals
- flowsheet modelling and optimisation
- life cycle analysis of the biorefining products.

The supporting industrial companies were Neste Oil Oyj, Metso Power Oy, Pöyry Engineering Oy and Vapo Oy. In the BioScen Steering group the companies were represented by the following persons: Kari I. Keskinen and Juha Lehtonen (Neste Oil), Jani Lehto (Metso Power), Antti Hasanen (Pöyry) and Kauko Isomöttönen (VAPO).

M.Sc. Juha Leppävuori worked as a visiting scientist in USA at Carnegie-Mellon University (supervised by prof. Lawrence Biegler) from February 2010 to March 2011, focusing on the optimisation of fermentation model parameters.

1.2 Project scope

During the last decade, utilisation of sustainable biomass and biotransformation processes has gained an ever increasing interest. The number of commercialised processes is rapidly increasing in production of chemicals and biofuels. Figure 1.1 illustrates this trend as global planned capacity of biobased chemicals. In Finland

numerous industrial R&D projects have been initiated during the last five years, mainly focused on production of biofuels but also on other chemicals and polymers. Tekes's BioRefine program has been running since 2007 and as a part of it, ForestCluster Ltd's Future Biorefinery (FuBio) was launched in 2009. Thus, empirical knowledge on processes using biomass is rapidly increasing, while less attention has been paid on their quantitative engineering and modelling.

The biorefinery uses biomass as raw material for production of fuels, hydrogen and chemicals. The production concept may be independent or integrated to a traditional manufacturing process, in particular to one of forest, food or forage industries with conventional products. The concepts for future biorefineries are often characterized according to their select of raw materials: ligno-cellulosic for the use of forest biomass and whole-crop for agricultural raw materials. A third alternative is the green biorefinery, by which basically fresh 'wet biomass' (grass, clover, algae etc.) are used as starting materials. Alternatively the technology platform – low-temperature fermentation or thermal treatment – is used as the classifying basis. Regardless of the chosen raw-material, the biorefining concept aims to produce either fuels, chemicals, material products or feed with a combination of different technologies. Two basic technology platforms are being used: hydrolysis of sugars followed by their bioprocessing to chemicals or thermal treatments in combination with subsequent chemical conversions to products.

The concept 'biorefinery' has inspired itself from the respective 'oil refinery' where mineral oil is used as raw material. It is well-known that petrochemical industry receiving its feedstock from oil refining produces quite limited number of basic chemicals, from which then a multitude of molecules are being further produced as chemical products. In this processing chain various phase separations and chemical conversions are applied.

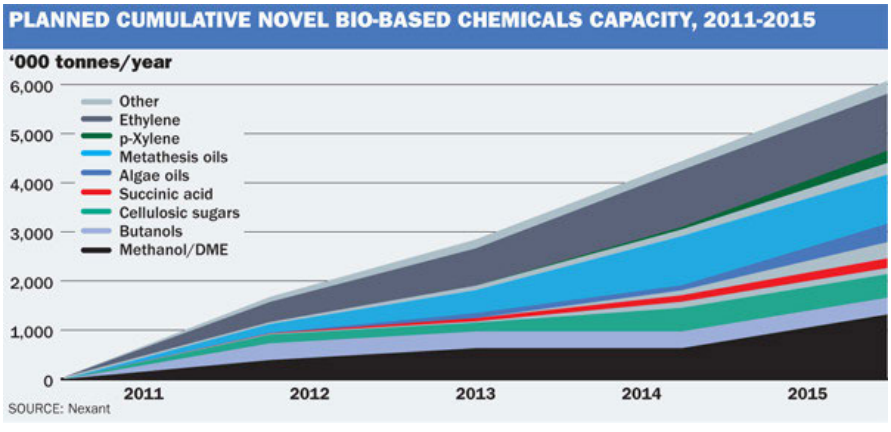

Figure 1.1. A recent analysis of planned bio-based chemicals production (*http://www.icis.com/Articles/2012/04/16/9549985/bio-based-chemicals-on-the-fast-track-to-commercialization.html*).

Oil refining has created the basis of chemical engineering. Its most important supporting science has been chemical thermodynamics. The design of all unit operations and processes will require the physical and chemical properties of components and mixtures. Unit operation models require quantitative thermodynamic models which are based on empirically determined thermodynamic data. Each stage of any planned new process is studied in terms of the methodology of chemical engineering science. As a result, it has become customary in chemical technology that no new processes are designed without constructing a simulation model. With advanced models, erroneous trials and costly experimental work can be minimized. The materialization of the biorefinery will also necessitate development of simulation methods and databases for biobased raw material and related unit processes. A parallelisation of the required methods is illustrated in Figure 1.2.

In general a simulation model precedes all new processes and targeted revamps in existing plants in chemical process technology. These successful engineering practices have also been increasingly applied in other fields. In Finland, particular attention has been in applying engineering models for pulp and papermaking as well as for other process industries.

In biorefining, there exist a number of different technology concepts ranging from high temperature gasification methods to aqueous fermentation based technologies and organic solutions. In many cases, the product is received in the form of a dilute aqueous solution with multitude of side products, requiring advanced fractionation methods. Thus, the requirements for process modelling also become more challenging.

With advanced modelling techniques using experimental data it is possible to evaluate reaction mechanisms and their kinetic rate parameters for complex biorefining systems. The pathways to final chemical products can involve a variety of routes from thermal processing to aqueous hydrolysis and metabolic reaction networks. As an alternative to chemical kinetic approach, thermodynamic equilibrium properties can often be efficiently used to screen both high temperature processes and biochemical pathways. Modern thermodynamic models will serve both improved understanding of various reaction networks including unit processes thereof and construction of reliable balance (flowsheet) models to evaluate the feasibility of industrial production units.

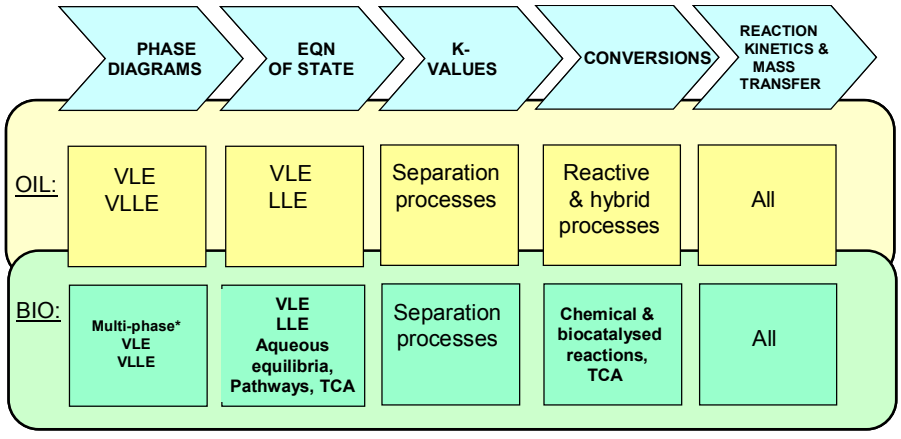

Figure 1.2. Comparison of the engineering models for oil refining and biorefining.

The BIOSCEN project was commenced within Tekes's BioRefine program as one of the few projects focusing on modelling and simulation of biorefinery processes. Thus a rather wide variety of methods were chosen with the aim to cover the versatility of potential applications. This text serves as a final report summarising the key contents and results of the project. Chapter 2 provides a brief description of the project and Chapters 3 describes the methods used in the project. Chapters 4 to 9 give a detailed description of the processes studied and key results. The topics covered in Chapters 4 to 9 cover the condensation stage of fast (biomass) pyrolysis, modelling of hydrodeoxygenation of pyrolysis oil, the thermal hydrolysis of lignocellulosic biomass for levulinic acid production, acid pretreatment of lignocellulosic biomass for its subsequent enzymatic hydrolysis and fermentation for bioethanol production and finally, o flowsheet modelling example for a bark refinery. Attention has been also laid on highlighting the current situation in the field and future challenges. The conclusions from the whole project are summarized in the Chapter 10.

2. Methods

This chapter outlines the physical property prediction and mathematical optimization, metamodelling and sensitivity analysis methods used in the BIOSCEN project. Creenhouse gas (GHG) calculation method used in isobutanol production analysis is also presented.

2.1 Physical property prediction

Fibre-biomass conversion into chemicals requires knowledge of involved components behavior. Variety of physico-chemical, reaction and transport properties are needed for the description of biomass conversion kinetics, distribution of compounds in different reactive zones and phase behavior of the components in different processes. Diversity and variability of the bio-products makes predictions of their properties a challenging task. However, most fibre-biomass processing methods represent techniques for biomaterial degradation. Some products of the degradation are well known compounds (like cresols, toluene, glucose), some of them have not been investigated earlier. A typical example of such degradation method is the pyrolysis of biomass for the production of 'biocrude', a mixture of liquefied components to be used as fuel or chemicals. Similarly the hydrolysis of biomass will lead to molecular fragments of the polymeric biomaterial, which then must be fractionated and recovered from the aqueous media.

Within the Bioscen project list of compounds was made that would supplement information needed for description of biomass pyrolysis process and hydrolysis of fibre-biomass Table (2.1 (a), (b), (c)). The lists are not complete, but serve to provide or verify information that has low quality or is not available for calculation of phase behaviour and kinetics in the commonly used process flowsheet simulator programs such as *Aspen Plus, Flowbat, ProII* etc. For description of pyrolysis product condensation one pseudo component was added to represent low molecular weight lignin in the system. The lignin is cut of from the poplar molecule structure (Boerjan *et al.*, 2003) to mimic average lignin molecular weight and elemental ratios (C:O:H) (Oasmaa *et al.*, 2003). Gaseous compounds are also taken into account because they influence equilibrium between vapor and liquids.

Table 2.1. Pyrolysis oil compounds. (a) pyrolysis product compounds; (b) pyrolysis reaction compounds (c) metabolites of hydrolysis and fermentation processes.

(a)	Name	Formula	CAS number		(b)	Formula	CAS number
1	water	H_2O	7732-18-5	1	water		
2	methanol	CH_4O	67-56-1	2	methanol		
3	formic acid	CH_2O_2	64-18-6	3	formic acid		
4	acetic acid	$C_2H_4O_2$	64-19-7	4	acetic acid		
5	guaiacol	$C_7H_8O_2$	90-05-1	5	guaiacol		
6	hydroxyacet-aldehyde	$C_2H_4O_2$	141-46-8	6	methane	CH_4	74-82-8
7	acetol	$C_3H_6O_2$	116-09-6	7	1,2-ethanediol	$C_2H_6O_2$	107-21-1
8	furfural	$C_5H_4O_2$	98-01-1	8	phenol	C_6H_6O	108-95-2
9	levoglucosan	$C_6H_{10}O_5$	498-07-7	9	toluene	C_7H_8	108-88-3
10	glucose-alfa D	$C_6H_{12}O_6$	492-62-6	10	ethanol	C_2H_6O	64-17-5
11	4-methyl-guaiacol	$C_8H_{10}O_2$	93-51-6	11	ethyl acetate	$C_4H_8O_2$	141-78-6
12	vanillin	$C_8H_8O_3$	121-33-5	12	cyclohexene	C_6H_{10}	110-83-8
13	abietic acid	$C_{20}H_{30}O_2$	514-10-3	13	cyclohexane	C_6H_{12}	110-82-7
14	low molecular weight lignin	$C_{22}H_{30}O_7$	-	14	benzene	C_6H_6	71-43-2
15	nitrogen	N_2	7727-37-9	15	catechol	$C_6H_6O_2$	120-80-9
16	carbon monoxide	CO	630-08-0	16	methylcyclo-hexane	C_7H_{14}	108-87-2
17	carbon dioxide	CO_2	124-38-9	17	benzaldehyde	C_7H_6O	100-52-7
18	methane	CH_4	74-82-8	18	o-cresol	C_7H_8O	95-48-7
19	ethane	C_2H_6	74-84-0	19	p-cresol	C_7H_8O	106-44-5
20	propane	C_3H_8	74-98-6	20	m-cresol	C_7H_8O	108-39-4
21	n-butane	C_4H_{10}	106-97-8	21	2,4-dimethyl-phenol	$C_8H_{10}O$	105-67-9
				22	hydrogen	H_2	1333-74-0
				23	carbon dioxide	CO_2	124-38-9
(c)							
1	pyruvic acid	$C_3H_4O_3$	127-17-3	5	alpha-acetolactic acid	$C_5H_8O_4$	918-44-5
2	2-butanol	$C_4H_{10}O$	78-92-2	6	alpha,beta-Dihydroxy-isovaleric acid	$C_5H_{10}O_4$	1756-18-9
3	isobutanol	$C_4H_{10}O$	78-83-1	7	2-ketoiso-valeric acid	$C_5H_8O_3$	759-05-7
4	isobutanal	C_4H_8O	78-84-2	8	3-hydroxy-3-methyl-2-oxobutanoic acid	$C_5H_8O_4$	6546-31-2
9-77	Furanose and pyranose forms of glucose, fructose and other sugar monomers; different forms of the sugar phosphates					$C_{5\text{-}6}O_{6\text{-}12}P_{0\text{-}2}$	-

The simulation of chemical processes requires knowledge of the physico-chemical properties of the treated compounds. The property factors in question may characterize the desired compounds from different perspectives, i.e. solvation behaviours of the compound, its activity in reactions, rate of mass transfer etc. Thus a selection

of the properties very much depends on the particular chemical process, through which the component originated. For example, pyrolysis of lignocellulosic biomass is one promising method for production of substitute of diesel fuels. Composition of the pyrolysis product is very complex (for example compared to respective gasification products) and represents a variety of bio-polymer derivatives and extractive compounds. The pyrolysis oil contains simultaneously polar, associations, polymeric compounds, carboxylic acids (i.e. formic and acetic) phenols, fused ring structures, lignin oligomers, alcohols and water. Moreover solvation of gases has a high importance both for pyrolysis oil stability (liquid phase splitting) and its production. Thus solubility, diffusion coefficients and variety of phase equilibria properties were selected to supply the needed information for pyrolysis oil production.

Another important field of biomass conversion is biological degradation. Most of these processes occur in water. Though properties of aquoeus solutions have been extensively investigated the quantitative description of biochemical reactions in water remains a challenging task. For example, the standard Gibbs energy of formation (G_f^0), which would give the respective equilibrium constants is unknown for many biologically active compounds. Therefore estimation of possible reaction routes for those molecules can not be performed directly and must be made using available G_f^0 estimation methods. However, accuracy of the currently available G_f^0 estimation techniques varies widely from 2 to 20 kJ/mol and can be lower for such kind of compound structures, which more rarely appear in conventional chemical engineering.

Within the BIOSCEN project the estimation of formation properties and solubility of bioactive compounds were made to facilitate development of models for fibre-biomass hydration and fermentation. Detailed description of properties estimations for pyrolysis oil reactions and condensation products, and for biomass fermentation methabolites are given below (see Section 3.1.3).

2.1.1 Phase equilibrium related properties

Description of any unit operation requires basic knowledge about the physical state of compounds (vapor, liquid, solid) and about the distribution of the component between different phases. The phase equilibria can be calculated in many different ways, but always applying the thermodynamic equilibrium concept, i.e. equal fugacity or chemical potential of the compound in all respective phases. A majority of the existing fugacity models employs critical properties of compounds and their vapour pressure as input parameters for the fugacity calculations. Thus critical properties and vapour pressure were estimated for all compounds from selected components list. These properties are shown in Table 2.2. The selection of a viable estimation method for unavailable properties was based on literature reviews, taking into account the size and chemical class of the molecules. Usually the most recent method of Nannoolal (Nannoolal *et al.*, 2004; Nannoolal *et al.*, 2008; Sanghvi & Yalkowsky, 2006) was used, with exception for small molecules where

Marrero – Pardillo method is more accurate (Marrero-Morejon & Pardillo-Fontdevila, 1999). In few cases, for which Nannoolal prediction deviates considerably from the other methods estimations, Joback or Stein method (Joback & Reid, 1987; Stein & Brown, 1994) were used because its reliability recommended e.g. by Poling (Poling *et al.*, 2001).

For the development of vapour pressure correlations (Pv) all available experimental information together with different estimation methods were compared and used. Coefficients of the vapour pressure correlation are given in Table 2.2.

Table 2.2. Critical properties of the pyrolysis oil compounds and their vapor pressures (PV).

	Tboil, K	**Tmelt K**	**Tc K**	**Pc MPa**	**Vc cm3/mol**	**PV at 50°C kPa**
Acetic acid	391.1e	289.8 e	592 e	5.8 e	177 e	7.58 e
Formic acid	373.7e	281.7 e	588 e	5.8pD1	125 pD2	17.28 e
Levo-glucosan	586pBv	386 e1	743pB1	5 pB2	365.5 pB1	7.21E-07 pB1
Glucose	691.1 pBv	419.2 e	846 pB1	5.2 pB3	490 pB3	1.01E-10 pB4
Acetol	418.6e	256.2 e	604.2 pB5	5.7 pB5	221.5 pB5	2.37 pBe
hydroxyacetaldehyde	423.2 e	358.2 e	629.7 pB5	6.5 pB1	167 pD	0.04 pBe
guaiacol	478.2 e	304.7 e	700.8 pB3	4.6 pB3	335.4 pB3	0.12 e
vanillin	558 e	355 e	800.6 pB3	4.6 pB3	402.4 pB3	7.53E-04 e
4-Methyl-guaiacol	494.2 e	278.7 pB6	710 pB3	4.1 pB3	402.1 pB3	0.07 pB3e
Abietic acid	649.7 pDv	446.7 e	865.1 pB3	2 pB3	1043.1 pB3	1.10E-06 pDe
methanol	337.9 e	175.5 e	512.5 e	8.1 e	117 e	55.53 e
water	373.2 e	273.2 e	647.1 e	22.1 e	55.9 e	12.35 e
furfural	434.9 e	236.7 e	670.2 e	5.7 e	252	1.34 e

*- vapor pressure correlation used for calculation the vapor pressure at 50°C,
e – experimental data, taken from DIPPR database; e1 – experimental data of Oja (Oja & Suuberg, 1999). **pD** – predicted in DIPPR project: e – fitted based on experimental data, v – from vapor pressure correlation, **1** – Ambrouse group contribution method (GCM), **2** – from family plot, **pB** – predicted within Bioscen project: e – fitted based on experimental data, v – from vapor pressure correlation, **1** – using Joback Stein GCM (Stein & Brown, 1994), **2** – using Joback GCM (Joback & Reid, 1987) with correction of Feng *et al*. (2005), **3** – using Nannoolal GCM (Nannoolal *et al.*, 2007), **4** – fitted to be between Nannoonal and Riedel GCM predictions taking into account Tb estimations of different methods, **5** – by Marrero-Morejon and Pardillo-Fontdevila GCM that is best for small molecule (Marrero-Morejon & Pardillo-Fontdevila, 1999), **6** – Marrero and Gani GCM (Marrero & Gani, 2001)

Table 2.3. Coefficient for vapor pressure correlations of pyrolysis oil compounds.

Name	Equation	A	B	C	D	E
acetic acid	1	39.4545	-6304.5	-4.30		6
formic acid	1	29.9911	-5131.030	-3.19		2
levoglucosan	2	9.3215	-13890.021	2.32		2
glucose-alfaD	3	15.4911	10749.922	-86.43		
acetol	3	9.3046	4608.641	-22.91		
hydroxyacetaldehyde	1	284.5054	-24639.248	-37.79		
guaiacol	1	237.7345	-17453	-33.72		2
vanillin	1	92.0145	-13646	-11.06		6
4-methyl-guaiacol	3	8.9182	4845.648	-61.84		
abietic acid	1	74.5935	-14612	-8.75		2
methanol	1	68.9025	-6904.5	-8.86		2
water	1	59.8335	-7258.2	-7.30		2
furfural	1	80.7545	-8372.1	-11.13	0.01	1
lignin fragment	3	14.2923	10753.149	-141.86		

*1 – ln(P/MPa)=A+B/T+C*lnT+d*TE*
*2 – ln(P/MPa)=A+B/T+C*lnT+d*TE*
3 – ln(P/MPa)=A-B/(T+C)

Other physical properties, such as heat capacities, formation energies, SRK (Soave-Redlich-Kwong) and Rackett equation parameters were estimated. Overview of the parameters sources is given in Table 2.4. The least investigated compounds were levoglucosan and 4-methyl-guaiacol. Some properties were not estimated due to unavailability or low accuracy of existing estimation technique for them.

Phase equilibrium properties are also very important for the description of pyrolysis processes in both reactor and condenser stages. Collection of available data, prediction of missing properties and modelling of the pyrolysis oil vapour-liquid equilibrium is described in Section 4.

Table 2.4. Collection of physical properties for pyrolysis oil compounds.

	Acetic acid	formic acid	levoglucosan	glucose-alfaD	acetol	Hydroxyacet-aldehyde	guaiacol	vanillin	4-methyl-guaiacol	abietic acid	furfural
Critical temperature											
Critical pressure											
Critical volume											
Normal boiling point											
Melting point											
Triple point temperature											
Triple point pressure											
Acentric factor											
Radius of gyration											
Solubility parameter											
Dipole moment											
Van der Waals reduced volume											
Vanderwaals Area											
Ideal gas heat of formation											
Ideal Gibbs energy of formation											
Ideal gas absolute entropy											
Heat of fusion											
Heat of vaporisation at normal boiling point											
Standard net heat of combustion											
Rackett parameter											
Surface tension at normal boiling point											
SRK acentric factor											
Molar volume at fixed T											
Density, Liquid											
Density, Solid											
Dynamic Viscosity, Liquid											
Dynamic Viscosity, Vapor											
Enthalpy of Vaporization											
Heat Capacity, Ideal Gas											
Heat Capacity, Liquid											
Heat Capacity, Solid											
Surface Tension											
Thermal Conductivity, Solid											
Thermal Conductivity, Liquid											
Thermal Conductivity, Vapor											
Vapor Pressure, Liquid											
Vapor Pressure, Solid											
Second Virial Coeff.											

– experimental data, – derived based on experimental data (Bioscen),
– predicted value (taken from literature), – predicted value (Bioscen)

2.1.2 Kinetics related properties

Formation energies of components are needed for prediction of reaction pathways of enzymatic hydrolysis and fermentation using energy diagrams (see Section 5 and 6). Partly the standard enthalpies and Gibbs energies of formation can be found in literature, but for many metabolites there is either no information available or the energy is provided for different conditions (other temperature or compound state). Supplementary information needed to translate all data into process conditions can yet be obtained by advanced prediction methods. The COSMO-RS model is a novel predictive technique that can be used for estimation of the solvation energies of gaseous and solid compounds based on molecular structure (Klamt, 2005). Fast screening of hydration energies for 68 metabolites for ethanol production were made using commercially available implementation of COSMO-RS model (COSMOtherm software). Procedure for the calculation is described in Section 4. The procedure usually includes establishing of the most stable structures of every compound under consideration and statistical thermodynamic calculations of solvation energies for those structures. A comprehensive search of most stable structures is a time limiting step in the solvation energy calculations, even though it improves accuracy of the energy estimation only a little. Due to large amount of compounds to be estimated such conformation search was not performed. However the final structure geometries were ones optimized starting from the structures available in common biomolecule databases (e.g. Beilstein database or NIST).

In addition, estimation of properties for iso-butanol production metabolites was performed; the results are shown in Table 2.5. Correlations for vapour pressures and enthalpy of vaporization were based on COSMO-RS model predictions. If the conformation analysis indicated that several structures are stable in liquid and vapour phases for these molecules, the conformers were taken into calculations. For solid state standard enthalpy and Gibbs energy prediction methods of Domalski (Domalski & Hearing, 1993) was used.

Table 2.5. Properties predicted for iso-butanol production pathways: Gibbs energy of hydration from vapor phase, standard solid state enthalpy of formation, vapor pressure at 30°C, enthalpy of vaporization at 30°C.

	N con-formers	G hydration (kJ/mol)	Hf_solid (kJ/mol)	Pressure at 30°C (kPa)	Hvap at 30°C (kJ/mol)
2ketoisovalericacid	6	-3.1	-649.33	0.00666	51303.13
3-hydroxy-3-methyl-2-oxobutanoic acid.mol	3	-10.4	-857.77	0.00059	63927.26
alpha,beta-Dihydroxy-isovaleric acid	10	-24.1	-963.86	0.00053	83165.76
alpha-acetolactic acid	9	-13.2	-893.5	0.04198	62583.96
ibutanal (2-methylpropanal)	2	8.5	-107.99	26.43490e	33483.39 e
ibutanol (2-methyl-1-propanol)	4	1.5	-336.8	1.98990 e	50687.13 e
pyruvic acid	4	-7.3	-592.76	0.23965 e	51631.91 e

e – values are taken from DIPPR database (DIPPR project 801).

2.1.3 Molecular transport properties

One of the most important transport properties of a compound is its diffusion coefficient. The diffusion coefficients were estimated for components of the pyrolysis oil (Table 2.2) in water and methanol. There are a number of different correlations for the estimation of unknown diffusion coefficients. However, classic ones such as Wilke-Chang (Wilke & Chang, 1955) give poor estimates for the diffusion coefficients of unideal components like oxygenates. Six different correlations for diffusion coefficients were compared: Stokes-Einstein, Wilke-Chang, Tyn-Calus, Sitaraman, Hayduk-Minhas and Nakanishi. Group contribution methods were also assessed, but found to have insufficient data available so far to be applicable. Of the different correlations, Tyn-Calus (Tyn, 1975) and Sitaraman (1963) methods are generally considered to be the most suitable for oxygenated components.

Tyn-Calus method is based on the use of the parachor, P, which describes the connection between molecular structure and physical properties. Parachor is an additive quantity and can be approximately expressed as a sum of empirical increments Pi corresponding to the single atoms or groups in the molecule. Parachors for many components can be found in the literature and it can be estimated using various methods. Tyn and Calus suggested a simple correlation for the parachor based on surface tension of a pure component:

$$P = V\gamma^{\frac{1}{4}}, \tag{2.1.1}$$

in which P is the parachor, V is the molar volume (cm^3mol^{-1}) and γ is the surface tension ($g\ s^{-2} = 10^{-3}\ N\ m^{-2}$). Using the parachor, binary diffusion coefficients can be calculated using the Tyn-Calus equation:

$$D_{AB}^{0} = 8{,}93 * 10^{-8} \frac{V_A^{\frac{1}{6}}}{V_B^{\frac{1}{3}}} \left(\frac{P_B}{P_A}\right)^{0,6} \frac{T}{\eta_B}, \quad (2.1.2)$$

in which D^0_{AB} is the diffusion coefficient of component A in solvent B, V_i is the molar volume (cm^3mol^{-1}) of the component i in temperature T (K), P_i is the parachor ($cm^3g^{1/4}s^{-1/2}mol^{-1}$) of the component i and η_B is the viscosity (cP) of the pure solvent B.

With good values for the parachor, the Tyn-Calus method should be good especially for estimating the diffusion coefficients of polar components. However there are some special considerations that have to be taken into account when using the Tyn-Calus correlation. Firstly it can not be used for solvents with high viscosities above 20 cP. Secondly, when water is the solute, it needs to be treated as a dimer, *i.e.* the values of its P and V need to be doubled. Thirdly, when the solute is an organic acid and the solvent is not water, methanol or butanol, it needs to be treated as a dimer. Fourthly, when the solute is non-polar and the solvent is an alcohol with a single hydroxyl group, the V and P of the solvent need to be multiplied with the association factor $n = 8*\eta_B$.

The Sitaraman correlation (Sitaraman, 1963) is based on using the heats of vaporization of the binary components. This makes the correlation widely applicable since the heat of vaporization is rather easy to obtain. Binary diffusion coefficient using the Sitaraman correlation:

$$D_{AB}^{0} = 5{,}4 * 10^{-8} \left(\frac{M_B^{\frac{1}{2}} L_B^{\frac{1}{3}} T}{\eta_B V_{A,b}^{0,5} L_A^{0,3}} \right)^{0,93}, \quad (2.1.3)$$

in which M_B is the molar mass ($g\ mol^{-1}$) of the solvent B, L_i is the heat of vaporization ($cal\ g^{-1}$) of the component i at its boiling point T_b, T is the temperature (K), η_B is the viscosity (cP) of the pure solvent B and $V_{A,b}$ is the molar volume (cm^3mol^{-1}) of the solute A at its boiling point T_b.

The Sitaraman correlation gives good estimations for systems with water as the solute and is usable over a wide range of conditions. However it does not account for the molecular interactions as well as the Tyn-Calus with well estimated values for the parachor.

Tyn-Calus and Sitaraman correlations were used to calculate the diffusion coefficients of the different binary pairs formed from the selected pyrolysis oil model compounds: water, methanol, acetic acid, formic acid, levoglucosan, glucose, acetol, hydroxyacetaldehyde, guaiacol, vanillin and 4-methyl-guaiacol. The values of the variables required by the equations were taken from the DIPPR 801 database. The temperature dependent variables were calculated using the DIPPR correlations presented as Equations (2.1.4)–(2.1.6) and the heats of vaporization required in the Sitaraman equation were converted from the enthalpy of vaporization calculated with Equation (2.1.6) using Equation (2.1.7):

$$\eta(\mathrm{Pas}) = e^{A+\frac{B}{T}+ClnT+DT^E}, \quad (2.1.4)$$

$$\gamma\left(\frac{\mathrm{N}}{\mathrm{m}}\right) = A(1-T_r)^{B+CT_r+DT_r^2+ET_r^3}, \tag{2.1.5}$$

$$\Delta H_{vap}\left(\frac{\mathrm{J}}{\mathrm{kmol}}\right) = A(1-T_r)^{B+CT_r+DT_r^2+ET_r^3}, \tag{2.1.6}$$

$$L\left(\frac{\mathrm{cal}}{\mathrm{g}}\right) = \frac{\frac{\Delta H_{vap,b}}{1000}}{\frac{M}{4{,}184\frac{\mathrm{cal}}{\mathrm{J}}}}, \tag{2.1.7}$$

The values of levoglucosan and 4-methyl-guaiacol, not found in DIPPR 801, were taken from Knovel Critical tables (Knovel critical tables, 2008) (molar mass), Yaws' Thermophysical Properties of Chemicals and Hydrocarbons (Yaws, 2010) (boiling point temperature) and predictions made by ACD/Labs (molar volume). As an example, the diffusion coefficients of the components in water and methanol calculated at 25 and 50°C are shown in Table 2.6.

Table 2.6. Binary diffusion coefficients for the model components in water and methanol according to Tyn-Calus and Sitaraman correlations.

Solvent	**Water**				**Methanol**			
	Tyn-Calus		**Sitaraman**		**Tyn-Calus**		**Sitaraman**	
Temperature	25°C	50°C	25°C	50°C	25°C	50°C	25°C	50°C
Component	DAB [*10-5 cm2s-1]							
Water					2.35	3.49	1.14E-11	1.93E-11
Methanol	1.41	2.48	1.14	2.04				
Acetic acid	1.26	2.24	1.29	2.30	2.39	3.53	2.31	3.28
Formic acid	1.35	2.36	1.49	2.65	2.55	3.72	2.66	3.78
Levoglucosan	N/A	N/A	1.15	2.05	N/A	N/A	2.06	2.93
Glucose	0.83	1.46	0.699	1.24	1.57	2.30	1.25	1.78
Acetol	1.09	1.93	1.07	1.91	2.06	3.04	1.91	2.72
Hydroxyacetaldehyde	1.14	2.02	1.17	2.08	2.14	3.18	2.09	2.97
Guaiacol	0.900	1.59	0.947	1.69	1.70	2.50	1.69	2.41
Vanillin	0.805	1.42	0.882	1.57	1.52	2.24	1.58	2.24
4-methyl-guaiacol	N/A	N/A	0.951	1.69	N/A	N/A	1.70	2.42

Unfortunately, estimation of the applicability of these two correlations for the estimation of the diffusion coefficients of the pyrolysis oil model compounds is difficult, since there is hardly any experimental data for the diffusion coefficients of most of these components. However, some differences become apparent when the values are compared over the whole temperature range where the equations can be applied. This can be seen in Figure 2.1, which shows the calculated values for the diffusion coefficients of water and guaiacol in water. The correlations agree fairly well except for guaiacol at temperatures above 600°C.

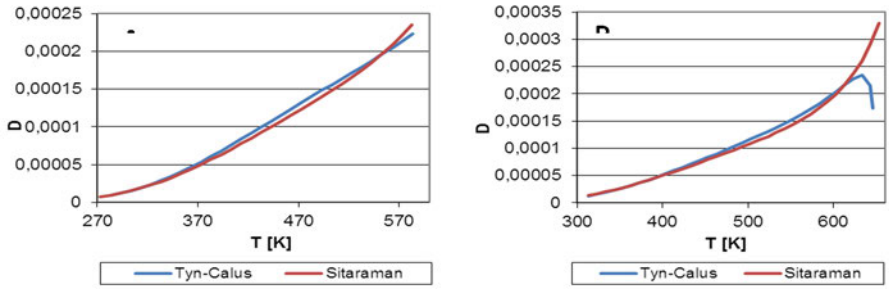

Figure 2.1. The diffusion coefficients of A) formic acid and B) guaiacol in water.

Many of the values of the variables used by the correlations are based on more correlations or calculations. Therefore, it seems a promising approach to model all molecular properties such as molecular volumes and the parachor with the same modeling tool in order to make all estimation of equal accuracy.

2.2 Sensitivity analysis and metamodels

Sensitivity Analysis (SA) will show in which measure the uncertainty in the output of a model (numerical or otherwise) can be apportioned to different sources of uncertainty in the model input. Two main approaches of the sensitivity analysis (SA) are local and global. Local SA is based on derivatives of the input/output variables and gives direct information on how output is affected as the input changes. As a drawback, derivatives are only informative at the base point where they are computed. Global SA, instead, strives to provide information covering the whole input space. It is based on the consideration that a handful of data points judiciously thrown into the design space is far more effective (with regard to informativeness and robustness) than estimating derivatives at a single data point (usually in the centre of input space). In the rest of this section, the focus will be on global SA.

General steps of GSA can be given as:

1. Establish the goal of analysis and form the output function (model) that answers the questions posed.
2. Decide which input factors are included in the model.
3. Choose a distribution function for each of the input factors (use literature, expert opinion, etc.).
4. Choose a proper sensitivity analysis method.
5. Generate the input sample according to chosen analysis method.
6. Evaluate the model using the generated sample.
7. Analyse the model outputs and draw the respective conclusions. If necessary, go back to Step 1.

To execute sensitivity analysis, there exist several software packages, and some of them are freely available. *SimLab*, is a free development framework for Sensitivity and Uncertainty Analysis. SimLab license encourages free non-commercial use and it provides a reference implementation of the various available uncertainty and sensitivity analysis techniques. The development and maintenance of SimLab is financed by the Unit of Econometrics and Applied Statistics of the Joint Research Centre (European Commission). More information about SimLab is available at http://simlab.jrc.ec.europa.eu/.

Another example of free software is *GUI-HDMR*, which is available as a *Matlab* toolbox with a graphical user interface. It provides a straightforward and efficient approach to explore the input-output mapping of a complex model with a large number of input parameters. In *GUI-HDMR* variance based sensitivity indices can be determined in an automatic way in order to rank the importance of input parameters and to explore the influence of parameter interactions. *GUI-HDMR* information and download is available at http://www.gui-hdmr.de/index.html.

2.2.1 Sensitivity analysis

In BioScen project, both *SimLab* and *GUI-HDMR* software were utilized, and they gave results of similar fashion. In BioScen project, we had one particular difference to most of the other SA literature. Instead of using mere *input factors* versus *output(s)*, we made further division of input factors in to *variables* and *parameters*. With this categorization parameters are something whose values should be fixed during the development of the model, and variables are something that the end user of the model uses to change, for example, the state of the model. This division was very natural, as BioScen project dealt with model development, and SA was used to support this.

Division to parameters and variables lead to a question of how should we treat them in sensitivity analysis, i.e. are they both considered as input factors in a regular way, or something else. Soon an obvious idea surfaced, and we defined our own way of global SA by studying whole region in variable space, and in each sample point in that space, each of the parameters was disturbed in turn (one-at-a-time) by predefined amount, and effect of this disturbance to output(s) was recorded. This scheme is henceforth referred to as the *global-local* scheme. Finally, several statistics for each of the parameters were calculated, such as mean and median effect on output. By these numbers it is easy to see which of the parameters has he most pronounced effect on the output, and on the other hand, it is easy to see which parameters have only negligible effect, and they may be left with minor consideration or even prune from the model.

In the Bioscen project several problems were studied, such as production of levulinic acid, yeast cultivation and fermentation, hydrolysis of biomass, and bark refinery. Especially several versions of the hydrolysis model were analyzed, and sensitivity analysis could be used for the development of the model. The SA with divisioned input factors remained our primary GSA tool as the latter half of the

project. Finally, as a comparison, a sensitivity analysis on one problem, **the bark biorefinery**, was run by treating variables and parameters on an equal footing.

2.2.2 Metamodelling

In metamodeling (also known as surrogate modeling and response surfaces), the original model is replaced with another model, *metamodel*, which should depict the original model as well as possible. Usually, the reason to this is that the original model is computationally expensive, and thus it may pose a prohibitive cost if optimization, sensitivity analysis or some other action requiring a high number of model runs is required. After a metamodel is generated based on input-output (training) data set, it is very fast to run it. Supposing that the modelling of a biorefinery would encounter exceptionally complex engineering simulations, the option of metamodels was studied as part of the BIOSCEN project.

Technically metamodels may be constructed in several different ways. Probably the four most commonly used techniques are artificial neural networks (ANN) (Haykin, 1998), kriging (Cressie, 1990), radial basis functions (RBF) (Buhmann, 2003) and support vector machines (SVM) (Cortes & Vapnik, 1995; Vapnik, 1998). They differ in principle by their internal working methods and mathematical formulations. In terms of training time, RBF and SVM are probably the fastest ones, naturally depending on the quality of the implementation. Another consideration is connection of metamodel to training data, i.e. RBF and kriging metamodels need to have training data incorporated in to them, which may be a problem in case of a very big training data set. On the other hand, ANN and SVM metamodels do not need a training data set, and in a sense we may postulate that they are very compact representations of the original model. We may illustrate this behaviour by assuming that output happens to be completely flat, and we are sampling it with a million points. RBF and kriging models need to carry all those million sample points along them, whereas ANN and SVM models end up being very compact in this case.

At the beginning of the BioScen project some comparison between different metamodeling techniques was executed. Based on these results, RBF models were selected to be used. As the topical metamodelling problems the biomass hydrolysis, the bioethanol production by fermentation and the simulated moving bed process for product fractionation were selected.

Results were mixed to some extent, in some output functions average error was very small (<<1%), and in some few percents. Although these numbers as themselves are not too disconcerting, it is worth to mention that maximum errors were in the range of tens of percents, although number of these is very low. It is up to the modeler to decide whether such uncertainty is acceptable or not.

In Figures 2.2–2.3 two different ways to display the quality of a metamodel are presented.

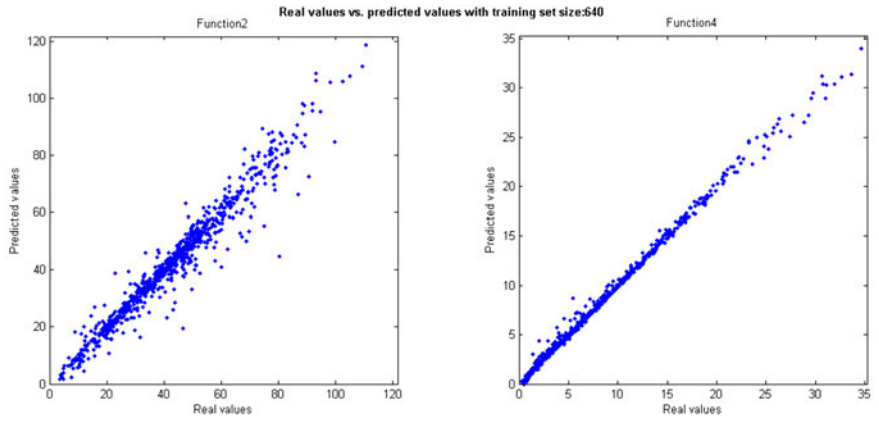

Figure 2.2. Quality of two output functions of metamodeled hydrolysis process. From this figure one can see what kind of values a metamodel has produced in contrast to the real model. If the metamodel accuracy was perfect, the line would be a thin line without any scatter around it.

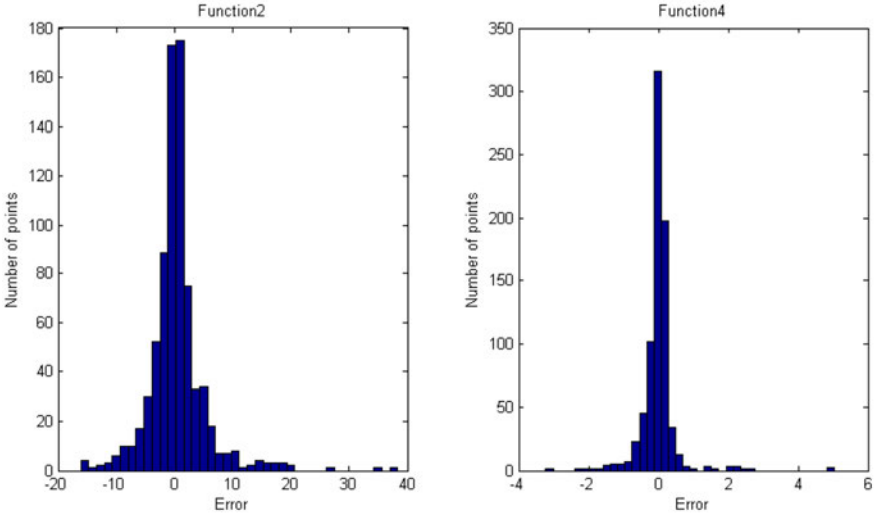

Figure 2.3. Histograms of errors of two output functions of metamodeled hydrolysis process. There are more occurrences of small errors for function 4, whereas function 2 has also bigger deviations. This same behaviour is clearly seen in Figure 4.

2.3 Optimization

There exist several different disciplines in optimization. Probably the most coarse useful classification of methods is in to a local, global and multiobjective optimization.

Local optimization algorithms (see e.g. Bazaraa et al., 1993; Nocedal & Wright, 2000) can only find the minimum which can be found only using either an ascending (maximization) or a descending (minimization) path. In this sense, local algorithms are trapped in to first objective function valley they find.

On the contrary, global optimization algorithms (see e.g. Törn & Zilinskas, 1989; Weise, 2009) strive to find the deepest of all valleys within a given search space. Multiobjective optimization (see e.g. Branke *et al.,* 2008) is yet wholly different, as then one must consider several objective functions at the same time, and their values are usually conflicting in a way that it is impossible to improve one value without degrading the value of at least one other objective. Also the concept of optimality is no longer straightforward, while it is defined by set of vectors containing objective function values, instead of one single value and there can be an infinite number of optimal solutions.

Currently, different evolutionary approaches are favoured among engineering disciplines. This is probably due to their ease of implementation and utilization, as well as availability of source code, and their ability to tackle several different kinds of problems. On the downside it must be mentioned that usually they require rather high number of objective function evaluations, thus rendering themselves almost useless in the case of computationally very demanding problems.

2.3.1 Optimization in BIOSCEN

In the BioScen project, optimization was combined with metamodelling techniques and then applied to the simulated moving bed separation process, originally presented by Kawajiri & Biegler, 2006. Simulated Moving Bed (SMB) processes are periodic adsorption processes for separation of chemical products based on liquid chromatography. They are applied to many important separations in sugar, petrochemical, and pharmaceutical industries and can be applied also in biorefineries. SMB processes are dynamic processes operating on periodic cycles which makes them challenging from the optimization point of view. Previously, SMB processes for separation of fructose and glucose have been optimized with respect to single (Kawajiri & Biegler, 2006) and with multiple objectives (Hakanen *et al.,* 2007) using gradient-based local optimization techniques. In this work the idea was to apply metamodelling techniques in optimization of the same SMB process in order to enable using global optimization techniques. Global optimization techniques have higher computational demand than local ones but they are more likely to produce better solutions for problems having multiple local optima.

The SMB problem has four objective functions: maximize throughput, minimize desorbent consumption, maximize purity of the product and maximize recovery of the product. The design variables included the liquid velocity in four different zones of the process as well as a step time which determines how often the place of the inlet (feed and desorbent) and outlet (extract and raffinate) streams are switched. Thus, altogether there were four objective functions and five decision variables. Details for modelling the SMB problem can be found in (Hakanen *et al.,* 2007).

The SMB process was metamodelled by using RBFs. The training data consisted of a Hammersley (1960) sequence of 2500 points in the five dimensional design variable space generated by using the SMB process model. The training of the RBFs took about five seconds for each objective function. It was found that throughput and desorbent consumption could be very well metamodelled (error << 1%) while purity and recovery were a little bit more difficult (errors were 2.6% and 6.0%, respectively).

The obtained metamodels were utilized in multiobjective optimization of the SMB process using global optimization methods and compared to previous results published in (Hakanen *et al.,* 2007). The results indicated that using the metamodelled objective functions, we could obtain better values for the objectives when compared to previous results with local optimization methods. In addition, the time used for generating one optimal solution for the multiobjective SMB problem was reduced from 1–2 minutes to about 15 seconds when metamodels and global optimization method was used instead of local optimization method and the SMB process model. On the other hand, when the solutions found by using metamodels were evaluated with the SMB process model they turned out to be slightly worse when compared to the solutions found by local optimization method and the SMB process model. The reason for this is due to the metamodelling difficulty for purity and recovery objective functions, although the benefit of finding almost as good solutions much faster still remains.

These findings were presented at the MCDM2011, the 21st International Conference on Multiple Criteria Decision Making held in Jyväskylä, Finland, June 13–17, 2011. In addition, an alternative approach to metamodelling and metamodelling-based optimization of an SMB process was made in (Giri *et al.*, 2011) where genetic programming (Collet, 2007) was used in metamodelling and an evolutionary approach was used in multiobjective optimization. This was a collaboration with professor Nirupam Chakraborti (Indian Institute of Technology, Kharagpur). It was initiated based on our presentation in the MCDM2011 conference and it offered an alternative approach to metamodelling of the SMB process. The findings of that study were quite similar and an article (Giri *et al.*, 2011) based on those was submitted to Computers & Chemical Engineering journal.

2.4 Methodology for greenhouse gas emission calculation

The European Union Directive on the promotion of the use of energy from renewable sources (RED) was published in June 2009 (EC, 2009). It establishes a mandatory target to increase the use of renewable energy sources in final energy consumption to a level of 20%, and in transportation to a level of 10%, by 2020 within the EU. For Finland, the differentiated target for renewable energy sources in final energy consumption is 38.0% in 2020, whilst the corresponding figure in the reference year 2005 was 28.5%. In addition, the RED introduced sustainability criteria for transportation biofuels and other bioliquids, setting requirements for the origin of raw materials and greenhouse gas (GHG) emission reduction compared

to reference fuels. The emission reduction should be at least 35% for biofuels and other bioliquids produced before the end of 2016. From the beginning of 2017, the target increases to 50% and from the beginning of 2018 to 60% for biofuel production installations where production begins after 1 January 2017. The sustainability criteria need to be met in order for biofuels and bioliquids to be accounted for renewable energy targets and to benefit from subsidies.

The RED provides default values for GHG emission reductions (%) compared to reference fossil fuels and disaggregated default values, separately and as an aggregate, for GHG emissions from cultivation, fuel processing and, transport and distribution for a range of biofuels. Liable actors may use these default values under certain conditions. Otherwise, they need to show that the actual GHG emissions value of their product is low enough to meet the given GHG emission reduction targets. In addition, the biofuel producer may always use the actual value instead of the default value. As a part of the sustainability criteria, the RED also introduces the first ever mandated method to calculate the actual GHG emissions of transportation biofuels and other bioliquids as well as the GHG emission reduction compared to fossil fuels.

The greenhouse gas emissions of ethanol and butanol were calculated following the method provided in the Renewable Energy Directive (RED) (EC, 2009).

The method is presented in the Annex V of the RED as follows:

$$E = e_{ec} + e_l + e_p + e_{td} + e_u - e_{sca} - e_{ccs} - e_{ccr} - e_{ee}, \quad (2.4.1)$$

where

E = total emissions from the use of the fuel

e_{ec} = emissions from the extraction or cultivation of raw materials

e_l = annualised emissions from carbon stock changes caused by land use change

e_p = emissions from processing

e_{td} = emissions from transport and distribution

e_u = emissions from the fuel in use

e_{sca} = emission saving from soil carbon accumulation via improved agricultural management

e_{ccs} = emission saving from carbon capture and geological storage

e_{ccr} = emission saving from carbon capture and replacement

e_{ee} = emission saving from excess electricity from cogeneration.

The emission from the manufacture of machinery and equipment are not taken into account. For further treatment, not all the parameters presented in the Equation 1 were assumed to be relevant. First, in the Commission decision on guidelines for the calculation of land carbon stocks (2010/335/EC) land use change is

referred as a change from one to another IPCC's land use category (forest land, cropland, grassland, wetlands, settlements, and other land) or seventh category that includes perennial crops and forest plantation. When straw is collected from the field, cropland remains unaltered, and it was assumed that no carbon stock changes due to land use change (e_l) occur. Second, according to the RED, the emissions from the fuel in use (e_u) equals zero in the case of bioliquids and biofuels (EC, 2009). Third, soil carbon accumulation (e_{sca}) via improved agricultural management was not assumed to take place when ethanol and butanol is produced from barley straw. Fourth, it was assumed that no carbon capture and storage system (e_{ccs} and e_{ccr}) is integrated to the processes. The Equation 1 reduces to the following form:

$$E = e_{ec} + e_p + e_{td} - e_{ee} \quad (2.4.2)$$

2.4.1 System boundaries

The system boundaries (Figure 2.4) were set in accordance with the framework given in the RED. Accordingly, wastes, i.e. agricultural residues, including straw shall be considered to have zero life-cycle emissions up to the process of collection of those materials. The RED determines that the unit of analysis for the purposes of allocation is the refinery if the fuels are produced in a refinery (EC, 2009). However, according to the particular definition it is unclear whether the ethanol/butanol process and the CHP plant should be assumed to be two separate units or one combined refinery. Therefore, both options were considered. Case 1 covers two separate units, and case 2 one combined refinery. Both the cases follow the principles of setting system boundaries as defined in the RED.

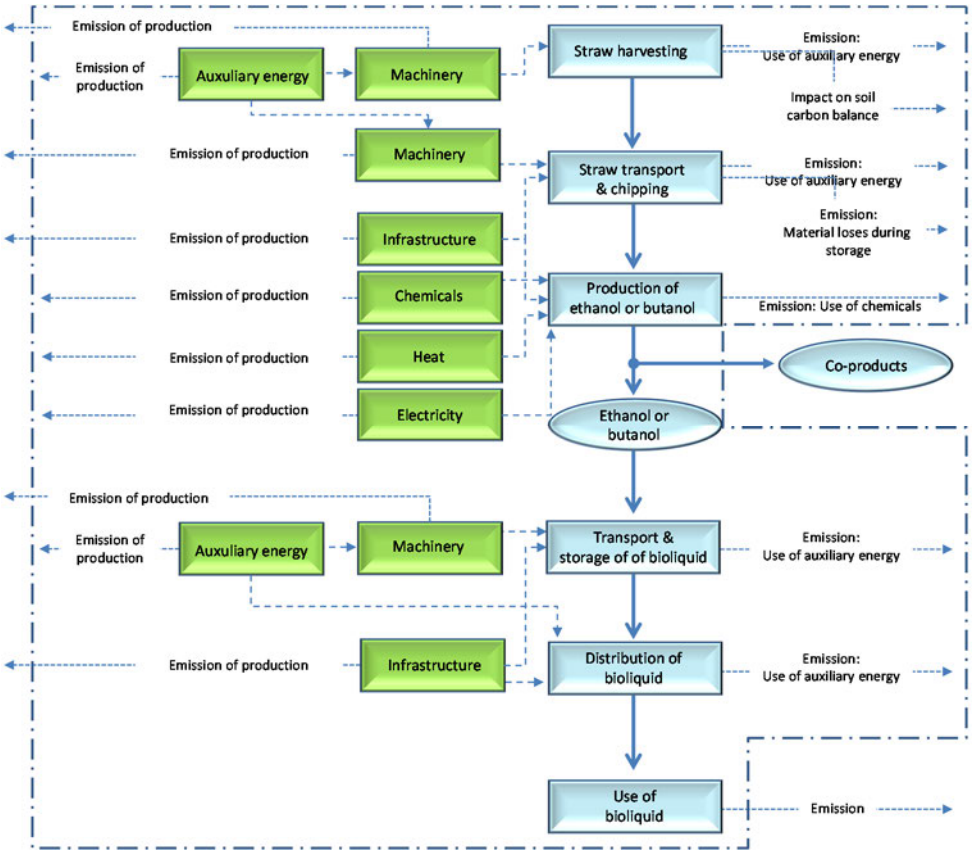

Figure 2.4. The system boundary according to the RED.

In case 1 (Figure 2.5) steam needed in the processing was assumed to be produced by combusting the ligneous-rich co-products of the process. The electricity needed in the processing was assumed to be purchased from the grid. In case 2, (Figure 2.6 both the steam and electricity were assumed to be produced using the ligneous-rich co-products of the process. Case 1 and case 2 leads to different methodological choices and therefore the results are expected to be different, too.

In case 1 the inputs to the ethanol/butanol process are raw material, electricity, steam, enzymes, yeast, sulphur acid and sodium hydroxide. Emissions from these are allocated between ethanol/butanol, co-product 1, and co-product 2. Emissions from production of steam consist of emissions from combustion of co-products which are allocated between electricity and steam. The calculations were simplified by omitting the feedback loop of emissions between ethanol/butanol process and the CHP plant.

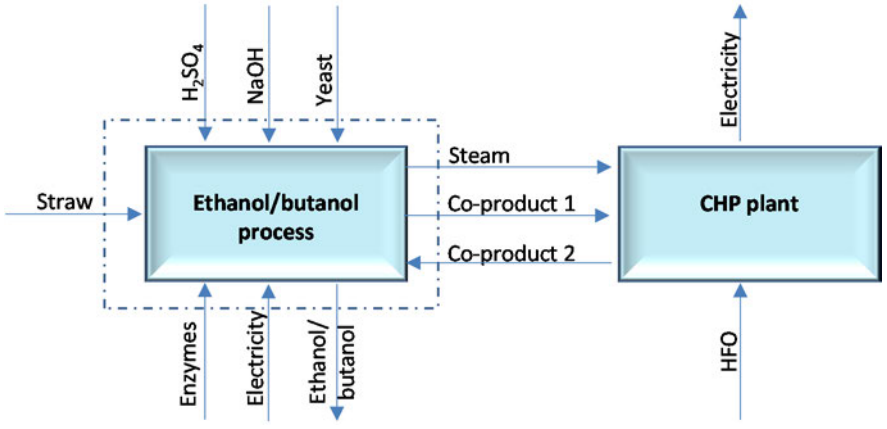

Figure 2.5. System boundary for case 1 whe re ethanol/butanol process and the CHP plant are considered as independent units.

In case 2, the inputs to the refinery are raw material, enzymes, yeast, sulphur acid, sodium hydroxide and, heavy fuel oil. Emissions from these are allocated between ethanol/butanol and, electricity. The electricity needed in the process was assumed to be taken from the CHP plant.

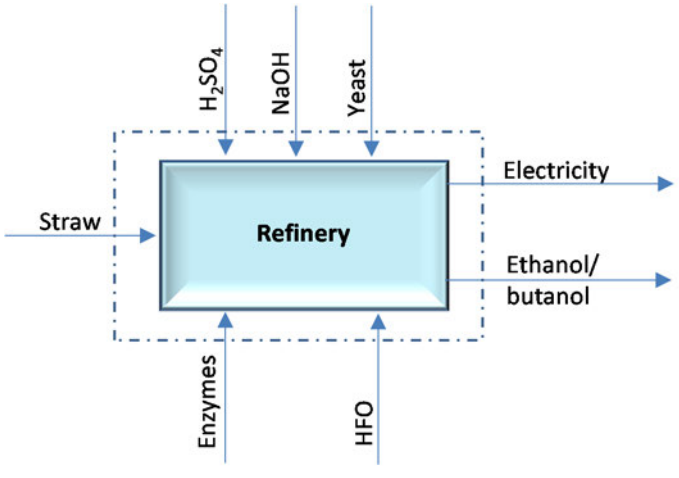

Figure 2.6. System boundary for case 2 whe re ethanol/butanol process and the CHP plant are considered as one combined unit.

2.4.2 Functional unit and allocation

The functional unit defines the quantification of the identified functions of the product. The purpose of the functional unit is to provide a reference to which the material and energy inputs and emissions of a system are related (EN ISO 14040, 2006). Identical functional unit is a must when comparing various product systems. According to the RED, the greenhouse gas emissions of fuels are expressed in terms of grams of CO_2 equivalent per MJ of fuel (EC, 2009). In this study, the difference on the GHG emissions of ethanol and butanol is expressed as per cents.

According to the RED, the allocation of emissions is based on the energy content of the products (EC, 2009). Energy content is determined by the lower heating value in the case of co-products other than electricity (EC, 2009). If a co-product has negative energy content, it is considered to have an energy content of zero in calculations (EC, 2009).

The greenhouse gases taken into account according to the RED are carbon dioxide (CO_2), methane (CH_4), and nitrous oxide (N_2O). In order to convert nitrous oxide and methane emissions to carbon dioxide equivalents, global warming potentials of 296 and 23 were applied in accordance with the RED.

3. Pyrolysis oil condensation

Concurrently, one of the most advanced biorefing technologies is the liquefaction of (forest) biomass by fast pyrolysis. After reaction at elevated temperature the pyrolysis products (components of the pyrolysis oil) are collected in a condenser stage. Condensation conditions are important for effective collection of the product, recycling gaseous reagents and for preventing the loss of organics. The pyrolysis oil is a very complex mixture of polar, associating, polymeric and aromatic compounds. Therefore the quantitative description of condensation of the pyrolysis oil is a demanding task.

3.1 Available equilibrium data

Extensive literature search was made for experimental equilibrium data that can be used for development of model for the condenser unit description. In Table 3.1 the available binary data are listed, in each cell of the table number of data set and total number of point (in brackets) are given. Only 15 experimentally measured binary systems were found from the 63 possible combinations of the selected 14 compounds for representation of pyrolysis oil (Table 2.4). No data were found for binary systems containing levoglucosan, hydroxyacetaldehyde, 4-methyl-guiacol, abietic acid and the lignin pseudo component. For description of those binaries either predictive models are needed or parameters of non-predictive models have to be estimated by other means. Additionally the condensation model should be able to take into account the carrier gas that serves to fluidize the reactor bed and thus will be present in the condenser in a large quantity.

Table 3.1. Binary data for selected pyrolysis oil compounds.

	Acetic acid	formic acid	levoglucosan	glucose	acetol	Hydroxyacetaldehyde	guaiacol	vanillin	4-methyl-guaiacol	abietic acid	methanol	water
formic acid	P; 5(74)											
levoglucosan												
glucose												
acetol												
hydroxyacet-aldehyde												
guaiacol												
vanillin												
4-methyl-guaiacol												
abietic acid												
methanol	P, T; 4(63)	T_artif; 2(58)					P; 1(10)					
water	P, T; 16(204)	P, T; 13(202)		P, T;S; 5(101)	T, P; 2(24)		P; 2(13)	S; 3(11)			P, T; 15(214)	
furfural	P; 2(30)	P*; 1(15)									HE; 1(10)	P, T, Az, 5+(63)

Where P – isobaric vapor liquid equilibrium (VLE) data, T – isothermal VLE data, Az – azeotropic data, S – solubility data, H^E – excess enthalpy data. T_art – data reconstructed from other binary and ternary data (the compounds are reactive), P* - data obtained by extrapolation of the experimental result to pure formic acid.

3.2 Model selection

Most important phenomena to be described in the pyrolysis condenser are vapour – liquid and liquid – liquid equilibria (VLE, LLE). Basic models rarely are able to describe vapour phase and liquid – liquid split simultaneously. One option is to use different models combination for separate description of vapour and liquid phases, so called gamma – phi approach (γ-ϕ) (Sandler, 2006). Additionally the same equation of states (EOS) with mixing rules based on excess Gibbs energy can be applied (ϕ - ϕ approach) as it is known to give acceptable accuracy of both phase description (Djordjević *et al.*, 2007). Classification scheme of models that can be applied to pyrolysis compounds phase behaviour are shown in Figure 3.1.

The gamma – phi approach will be considered first. Predictive models have to be used due to absence of 80% of data needed for description of VLE equilibria in

the system (see Table 3.1). On the other hand use of available experimental information is beneficial. We decided to employ UNIQUAC model, where missed binary interaction parameters are fitted based on COSMO-RS (Klamt, 1995) or COSMO-SAC (Lin & Sandler, 2002).

For the description of the vapour phase several options are available. In particular for condensation of the pyrolysis product in reactors where high rate of inert gas is used, ideal vapour phase description is a reasonable choice. For other cases EOS can be used. Peng Robinson (PR), Soave Redkich Kwong (SRK) or Mathias Copeman modification of SRK (MC- SRK) EOS are widely used. For the pyrolysis oil tested compositions no solution of VLE equilibrium was found when equations of state were combined with the UNIQUAC model, probably due to numerical problems (Flowbat software was used for the modelling (Keskinen & Aittamaa, 2010)). Description of the vapour phase using Henry's law approach requires fitting of the Henry's law constant based on experimental data for VLE behaviour of the compounds. Though the constants can be predicted by COSMO-RS model we decided to leave out this option due to many approximations and predictions that have to be used in this case both for vapour and for liquid phase. Nothnagel and Hayden O'Connell EOS (Hayden & O'Connell, 1975; Nothnagel *et al.*, 1973) can be used assuming that heavy components are presented in liquid phase only. However, those EOS do not provide sufficient means to describe the behaviour of heavy compounds correctly.

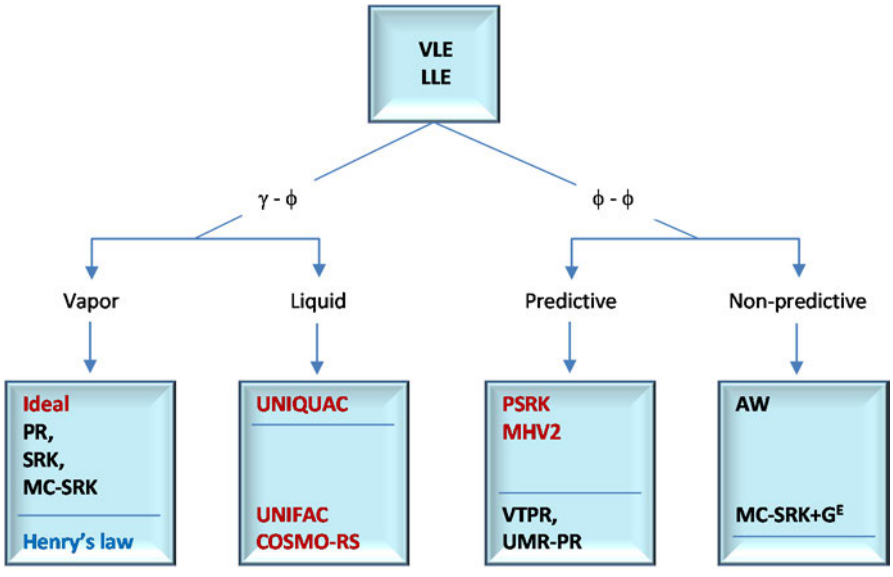

Figure 3.1. Scheme of models that can be applied for description of phase equilibria of pyrolysis oil components.

Recently modified EOS with mixing rules based on liquid phase excess enthalpy found to be applicable for description of very complex mixtures containing polar and asymmetric compounds (Djordjević *et al.,* 2007). Many of them are predictive (see Figure 3.1): PSRK (Holderbaum & Gmehling, 1991), MHV2 (Dahl *et al.*, 1992), VTRP (Ahlers & Gmehling, 2001), UMR-PR (Voutsas *et al.*, 2004)). Volume translated Peng Robinson (VTRP) and uniform mixing rule Peng Robinson (UMR-PR) are relatively new models where interaction parameters are estimated based on group contribution approach. Currently the parameters matrix is relatively small and therefore estimation of the pyrolysis liquid mixture VLE is not possible, mostly because of absence of the oxygen containing groups. Deficiency of parameter matrix is also a problem for another type of predictive model, i.e. Cubic Plus association (CPA, (Kontogeorgis *et al.*, 1996)). The model uses SRK equation of state with additional association parameters which are estimated based on group contributions.

Though pyrolysis compounds are not well investigated, non-predictive models for description vapour and liquid phases have some potential of use. In this work the binary data were simulated with COSMO-RS/SAC and parameters were fitted to UNIQUAC model. Therefore UNIQUAC can be used as an excess Gibbs energy (G^E) model to a selected equation of state, thus i.e. MC-SRK + MHV2 mixing rule based on UNIQUAC model combination was tested.

Another EOS+G^E type of model is Abovski-Watanasiri (AW) EOS (Abovsky & Watanasiri, 1998). The model was shown to be better then PSRK or MHV2 models (Abovsky & Watanasiri, 1999), but for fitting the mixing parameter such experimental information would be required, that is not available for the pyrolysis oil.

Statistical Associating Fluid Theory (SAFT, (Gross & Sadowski, 2000)) and it's extensions for polymeric and polar compound (perturbed chain PC-SAFT and p-PC-SAFT) is an additional option, often applied to many binary and many component systems, including the pyrolysis oils (Feng *et al.,* 2005). Developed in literature pure component and binary parameters of the SAFT can be used for similar compounds in multi-component mixture calculations. In this case, utilization of available parameters probably would not lead to significant differences in the VLE calculation results due to the moderate overall accuracy of the estimates.

Summarizing, the models presented in Figure 3.1 in green colour were tested for the pyrolysis oil description (see results below), the models that are shown in black can not be applied and the models in blue colour could be adjusted for the pyrolysis oil description, but additional work is needed for development of the required data or parameters. Testing of the last models will be continued within a project sponsored by Fortum Research Fund.

3.3 Parameters for modeling

For prediction of pyrolysis oil condensation two types of models were finally selected; ϕ- ϕ models – EOS+G^E (PSRK, MHV2) and γ – ϕ models. In the gamma – phi model excess Gibbs energy UNIQUAC model is applied for a liquid phase description and ideal and Hayden O'Connel models for a vapour phase. The se-

lected phi – phi type of models calculates G^E for liquid mixture based on UNIFAC or UNIFAC-Dortmund group contribution approaches and thus they are purely predictive in the G^E part. The only parameter that should be estimated for those models is an attraction parameter in EOS part (see Equation 3.3.1 for SRK EOS, *a* is an attraction parameter). Pure component vapour pressure calculated with this technique should correspond to experimental pure component pressure. This parameter can be estimated based on an acentric factor (ω), but in this case accuracy of the vapour prediction is low. For compounds selected as the pyrolysis oil representatives a pure component vapour pressure has been estimated by different method (see Section 4.1) and thus we recalculated attractive parameters for all compounds in form of Mathias Copeman Equation (3.3.1). The Matias Copeman correction of SRK EOS uses three parameters (c_1, c_2, c_3) for description of the attractive term of EOS. The parameters and the fitting error are given in Table 3.2.

$$P = \frac{RT}{v-b} - \frac{a(T)}{v(v+b)}, \quad a(T) = a(T_C)\alpha(T_R, c_1, c_2, c_2)$$
$$b = 0.08664\frac{RT_C}{P_C} \qquad (3.3.1)$$
$$a(T_C) = 0.42747\frac{R^2T_C^2}{P_C}$$
$$\alpha(T_R, c_1, c_2, c_2) = \left[1 + c_1\left(1-\sqrt{T_r}\right) + c_2\left(1-\sqrt{T_r}\right)^2 + c_3\left(1-\sqrt{T_r}\right)^3\right]$$

Table 3.2. Parameters of the Mathias – Copeman equation fitted based on pure component vapor pressure curves and accuracy of the fitting.

		c1	c2	c3	ΔP%* (average)
1	Acetic acid	1.4975	-2.8307	5.0355	1.3
2	Formic acid	1.0245	-0.3630	-0.4144	0.03
3	Levoglucosan	2.7614	1.9045	-8.1453	1.7
4	Glucose	4.6027	-10.4291	14.6724	1.7
5	Acetol	2.2569	-5.8903	8.6230	1.6
6	Hydroxyacetaldehyde	-0.9284	16.1439	-21.0200	1.0
7	Guaiacol	1.7359	-4.2727	8.5490	1.3
8	Vanillin	1.3298	0.6010	-0.7860	0.3
9	4-methyl-guaiacol	1.3140	0.2439	-1.2637	2.0
10	Abietic acid	1.0980	6.1236	-12.3805	3.1
11	Methanol	1.4381	-0.8554	0.5352	0.5
12	Water	1.1014	-0.7430	0.8546	0.2
13	Furfural	1.0362	0.0709	-0.5238	0.2
14	Lignin fragment	3.6704	-4.5856	6.3982	5

*$\Delta P\% = \mathrm{sum}(|P_{exp} - P_{meas}|)/P_{exp}$

Another type of modeling (γ-ϕ approach) requires parameters for liquid and vapour phase description. For ideal vapour assumption no parameter estimations are needed. For Hayden O'Connell EOS dipole moment, radius of gyration and group association parameters were obtained. Dipole moments and radii of gyration were either found in literature or estimated based on the molecules geometry with Quantum Chemistry software – Material Studio (Accelrys, 2007), DMol module was used for dipole moment calculation and Forcite module – for radii of gyration calculations. Group association parameters were set in accordance with molecular chemical class. All parameters are given in Table 3.3.

For the description of the liquid phase in the gamma – phi approach the selected model was UNIQUAC due to its ability to describe liquid – liquid equilibrium and the relative simplicity of the model. The model requires fitting binary interaction parameter τ_{ij} for description residual term of excess Gibbs energy (Sandler, 2006). The interaction parameter depends on interaction energy difference (Δu_{ij}) as it is shown in Equation 3.3.2. For most binary systems the interaction energy differences were described as temperature dependent (a_o and a_1 parameters were fitted) to increase accuracy of VLE prediction. The coefficients are given in Table 3.4. Note that the parameters are not symmetric, i.e. $a_{ij} \neq a_{ji}$. For system where experimental data were available, the a_0, a_1 parameters were fitted based only on the experimental data (green values in the Table 3.4, the components order is the same as in Table 3.3. For other binary mixtures the data for fitting were obtained with COSMO-RS/SAC model (see next section).

$$\tau_{ij} = \exp\left(-\frac{\Delta u_{ij}}{T}\right)$$
$$\Delta u_{ij} = a_{ij}^{0} + a_{ij}^{1} T \qquad (3.3.2)$$

Table 3.3. Parameters of Hayden O'Connell model for the pyrolysis oil compounds.

		Dipole moment	Radius of gyration	Association and solvation parameters												
		A	debye	1	2	3	4	5	6	7	8	9	10	11	12	13
1	Acetic acid	2.595	1.74	4.50	4.50	2.50	2.50	1.80	1.60	2.50	2.50	2. 50	4.50	2.50	2.50	1.60
2	Formic acid	1.48	1.52	4.50	4.50	2.50	2.50	1.80	1.60	2.50	2.50	2. 50	4.50	2.50	2.50	1.60
3	Levoglucosan	2.03c	4.839c	2.50	2.50	1.55	1.55	1.00	0.8	0	0	0	2.5	1.55	1.55	0.8
4	Glucose	4.75c	2.7491c	2.50	2.50	1.55	1.55	1.00	0.80	0	0	0	2.5	1.55	1.55	0.8
5	Acetol	2.98	4.287	1.80	1.80	1.00	1.00	0.90	1.00	1.00	1.00	1. 00	1.80	1.00	1.00	1.00
6	Hydroxyacetaldehyde	2.51c	2.3414c	1.60	1.60	0.8	0.80	1.00	0.58	0.80	0.80	0. 80	1.60	0.80	0.80	0.55
7	Guaiacol	4.02c	2.4583c	2.50	2.50	0	0	1.00	0.80	0.32	0.32	0. 32	2.50	1.55	1.55	0.80
8	Vanillin	4.7c	2.869c	2.50	2.50	0	0	1.00	0.80	0.32	0.32	0. 32	2.50	1.55	1.55	0.80
9	4-methyl-guaiacol	2.43c	1.014c	2.50	2.50	0	0	1.00	0.80	0.32	0.32	0. 32	2.50	1.55	1.55	0.80
10	Abietic acid	6.82c	2.2638	4.50	4.50	2.5	2.5	1.80	1.60	2.50	2.50	2.50	4.50	2.50	2.50	1.60
11	Methanol	1.536	1.71	2.50	2.50	1.55	1.55	1.00	0.80	1.55	1.55	1. 55	2.50	1.63	1.50	0.80
12	Water	0.615	1.83	2.50	2.50	1.55	1.55	1.00	0.80	1.55	1.55	1. 55	2.50	1.50	1.70	0.80
13	Furfural	3.17	2.3	1.60	1.60	0.8	0.8	1.00	0.55	0.80	0.80	0.80	1.60	0.80	0.80	0.58

c – values are calculated within Biocsen project by Accelrys Material Studio software. All other values are taken from literature (Prausnitz, *et al.* 1986.)

Table 3.4. UNIQUAC parameters values for the pyrolysis oil compounds.

a^0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1		75.32	173.6409	-33.9655	-694.094	-282.04	-60.8272	114.9209	-41.6828	53.37	171.94	32.7	-314.19	259.28
2	-136.21		257.45	-0.36424	-652.238	-937.131	4.517977	155.508	35.73191	143.38	-278.88	-40.62	-493.36	-441.259
3	-299.296	-591.805		-71.5211	-290.529	-76.0126	422.021	11.6181	414.395	298.2208	-5.45086	131.1319	2.34116	149.122
4	41.31879	-140.797	79.0667		-377.188	192.4011	-138.697	6.22656	-146.18	52.08486	-71.0042	-26.2906	72.0278	307.5697
5	416.103	225.3019	233.446	301.2963		-101.03	-561.913	-336.249	-540.518	238.9345	272.2417	787.782	21.57919	18.92071
6	380.141	451.372	66.1021	-163.971	154.779		554.756	110.624	447.6475	395.0288	187.6528	744.1968	33.2057	77.79713
7	242.6349	296.1212	-222.599	324.8596	445.1799	-292.126		-300.898	31.6123	-6.99435	14.01111	717.883	249.4886	-315.494
8	-171.126	-327.274	56.5398	109.941	228.345	-26.3227	226.56		197.585	162.8483	6.40846	687.864	-52.4882	-124.03
9	262.6647	286.1091	-222.13	358.2041	539.4355	-279.725	-20.7355	-260.676		-50.831	142.2862	505.4307	417.666	-281.302
10	151.92	265.04	-0.24793	171.4882	-303.837	-292.55	75.25328	-148.339	107.8016		-333.21	-126.12	-258.75	-251.866
11	-392.26	-417.71	24.60441	-63.7101	-67.3528	30.51897	71.64165	101.8994	42.96228	200.09		23.11	102.51	5.883666
12	119.2	-182.01	4.16874	25.91576	-504.542	-450.72	-200.688	-491.433	187.0114	526.24	14.77		383.32	372.94
13	172.77	119.6	68.2513	-0.63112	-0.24886	40.8303	-308.699	44.95909	-478.8	252.04	345.34	276.42		71.24908
14	-495.395	250.632	-90.7397	-155.411	-1.80395	21.38196	365.663	159.289	342.094	368.3912	134.7535	-256.385	21.9051	
a^1	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1		-0.1247	-0.57875	-0.23677	1.66147	0	-0.2629	-0.63837	-0.236	-0.32513	-0.2641	-0.04922	0.71219	-0.78834
2	0.2583		-0.6992	-0.18209	1.091395	2.22712	-0.22659	-0.51645	-0.23828	-0.3871	0.40031	0.47367	1.03755	0.526982
3	0.987448	1.64053		0	0.765119	0.340766	-0.6056	0.6279	-0.72275	-0.38444	0.53207	0.533946	0.65791	-0.28078

a^0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
4	0.495274	0.849127	0		0.858524	-0.13803	0.958013	0.765026	0.918005	0.229555	0.714577	0.989481	0.698095	-0.23428
5	-1.09605	-0.33649	-0.60977	-0.66587		-0.15245	1.23392	0.95603	1.14744	-0.70396	-0.17776	-1.60586	0.101715	-0.47021
6	0	-1.24174	-0.252	0.187329	0.143322		-0.59473	0.193892	-0.73653	-0.8972	-0.02491	-1.55699	-0.16403	-0.482
7	0.243419	0.127652	0.288439	-1.02625	-1.08469	0.178625		0.931789	0.287891	-0.10633	0.668828	-1.85529	-0.45816	0.572796
8	0.938066	1.082834	-0.61139	-0.75381	-0.78387	-0.3165	-0.72435		-0.70489	-0.46486	0.591275	-1.92168	0.108042	0.070406
9	0.148406	0.187744	0.424202	-1.04969	-1.19106	0.478345	-0.25378	0.917808		0.01987	0.458755	-1.08319	-0.99102	0.486311
10	0.47785	0.60587	-0.07921	-0.48279	0.960775	0.981457	0.081001	0.552431	-0.07542		1.63218	1.08063	0.99559	0.137601
11	0.63694	0.78179	-0.21286	-0.05986	0.002843	-0.22112	-0.4588	-0.49303	-0.35632	-0.60108		-0.05596	-0.53162	-0.26784
12	-0.21836	-0.12559	-0.36209	-0.47877	1.348345	1.195059	1.40226	1.97178	0.277117	-1.026	-0.01136		-0.76152	-1.25012
13	-0.44233	-0.20302	-0.61132	47.2711	-0.21659	-0.06568	0.602816	-0.06124	1.21569	-0.69292	-0.1557	-0.53654		-0.41477
14	1.62678	0.476894	0.190351	0.014583	0.638253	0.548612	-0.57888	-0.05673	-0.52716	-0.26182	0.70847	1.886834	0.531379	

3.4 COSMO-RS and COSMO-SAC models

In accordance with COSMO-RS model (Klamt, 2005) calculated screening charge distribution on molecular COSMO -surface is used for statistical thermodynamic treatment of interaction in the real solvent. Two modifications of the model implementation exist: one is made by COSMO-RS model author and usually referred as COSMO-RS (Klamt *et al.*, 2002), another was made in group of prof. Sandler and is called COSMO-SAC (Lin & Sandler, 2002). Both of them were used for comparative purposes.

The screening charge distribution and molecule geometry optimization are calculated with Density Functional Theory (DFT) applying Becke –Perdew (BP) functional with triple-zeta valence plus polarization basic set (TZVP). Turbomole program (version 6.0, (Schafer *et al.*, 2000)) was used for the COSMO-RS and the DMol module of Material Studio software (Accelrys, 2007) was used for the COSMO-SAC model. Two resulting screening charge distributions are slightly different and therefore it is recommended in literature to use the same DFT implementation that was employed for the models parameterization. Subsequent COSMO-RS thermodynamic calculations were done with the Cosmotherm program (COSMOtherm-C2.1, ("COSMOlogic GmbH & Co", 2009)) for COSMO-RS model and with open-source software for the COSMO-SAC model (COSMO-SAC, 2010; Mullins *et al.,* 2006).

Conformational search was needed at Quantum calculation stage for some molecules to find the most stable in solutions and in gas phase molecule structures. The search was made on lower then DFT computational level (Parameterized Model number 3 (PM3)) with HyperChem quantum mechanical software (Hypercube, 2002). Many conformers were found for levoglucosan, glucose and abietic acid molecules, i.e. 8, 9 and 7 lowest in energy conformers were taken for the molecules thermodynamic calculations. Phenol derivatives: guaiacol, vanillin, 4-methyl-guaiacol were found to have 2, 3 and 3 important conformers correspondingly. For small molecules, formic acid, acetol, hydroxyacetaldehyde, 1, 4 and 3 conformers are taken into account. Water, methanol, furfural and acetic acid input files for VLE calculations were taken from database provided together with the COSMOthem and from database created for the COSMO-SAC calculations (COSMO-SAC, 2010). Thus no quantum mechanical calculations for water, methanol, furfural and acetic acid were needed. Additionally formic acid, acetol, vanillin and abietic acid were presented in the VT-2005 Sigma Profile Database and they were taken without recalculation into COSMO-SAC model. In Figure 3.2 one can see the screening charge distribution for some selected pyrolysis molecules calculated in the Turbomole program.

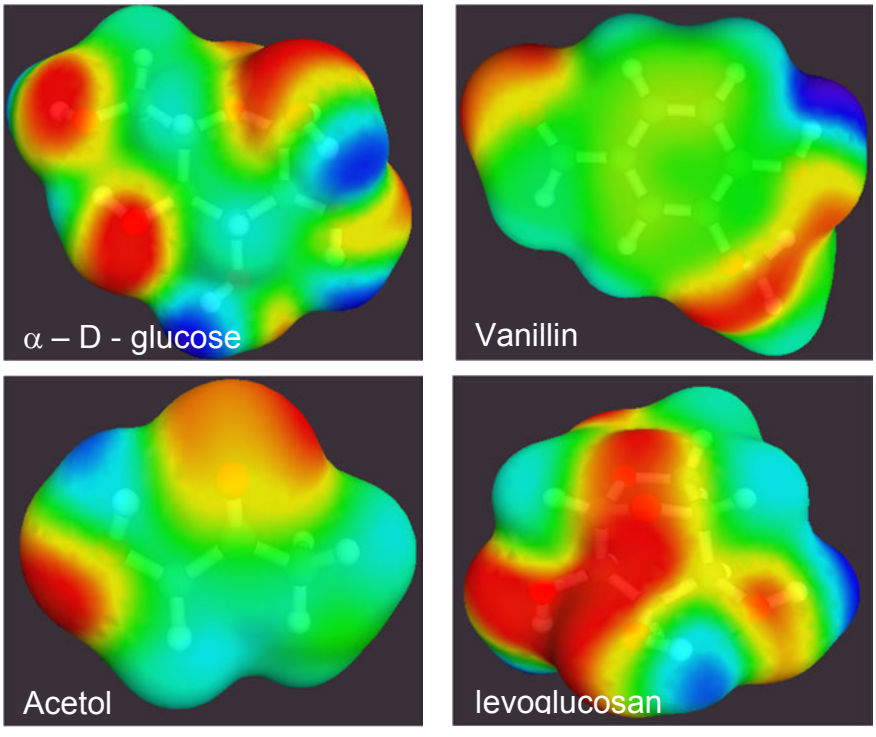

Figure 3.2. The screening charge distribution for glucose, vanillin, acetol and levoglucosan molecules.

Based on the calculated screening charge distribution, VLE data of the pyrolysis oil compounds were calculated for the binary mixture at temperature range from 25°C to 350°C. The data were used for UNIQUAC activity coefficients fitting. Replacement of COSMO-RS/SAC predictions with UNIQUAC model probably reduces accuracy of VLE estimation, but give us way to use available experimental properties together with the estimations. Moreover it speeds up calculations that could be important issue in description of multi-component multi-conformer flowsheet calculations.

3.5 Model verification

3.5.1 VTT experiment, other experimental data

Verification is an important stage for final selection of models for a process description. Experimental data for binary systems was used for deriving parameters of the models and thus could not be used for the models verification. There are very little data available

for ternary and multi-component system containing the pyrolysis oil compounds. The best way of verification would be to use existing pyrolysis reactor experimental data.

The experiments performed in VTT by Christian Lindfors (Lindfors, 2008) were decided to be used because it provides detailed information about condition of the experiments and analysis of the product. In the VTT experimental equipment the condensation stage of the pyrolysis process consist of several units: sieve tray condensation column, electric precipitator, spray condenser and intensive cooler. The last one serves to complete overall mass balance of the equipment. Important is that half of the reactor products come to electric precipitator, where aerosol particles are broken and condensed. Description of aerosol particle formation, breakage and organic loss related to aerosol flow was beyond the scope of our work and additional simplifications were taken to recalculate the pyrolysis condensation.

The condensation was modelled for the first unit (sieve tray column) with a flash VLE calculation and for products of two units (sieve tray column and electric precipitator) with two consequent flash calculations at sieve tray columns temperature and at room temperature correspondingly. Feed compositions of the condenser were not measured, but they were reconstructed from overall sum of products of condenser units including product gases that were not condensed but analyzed by a flow meter and gas chromatograph. In addition, a simplification of the product composition was adopted because in the experiments only fractional analysis of the pyrolysis product was performed, i.e. weights of fractions soluble in water, diethyl-ether and dichloromethane were measured. Such rough analysis does not provide information about precise composition of the pyrolysis oil and therefore approximate division of the pyrolysis oil selected compounds (Table 2.1(a)) between the fractions and following reconstruction of the product composition was made.

Predicted concentrations of the pyrolysis oil compounds that in our calculations serve as a lamp compound are shown in Figure 3.3 for first condenser unit operated at 47°C.

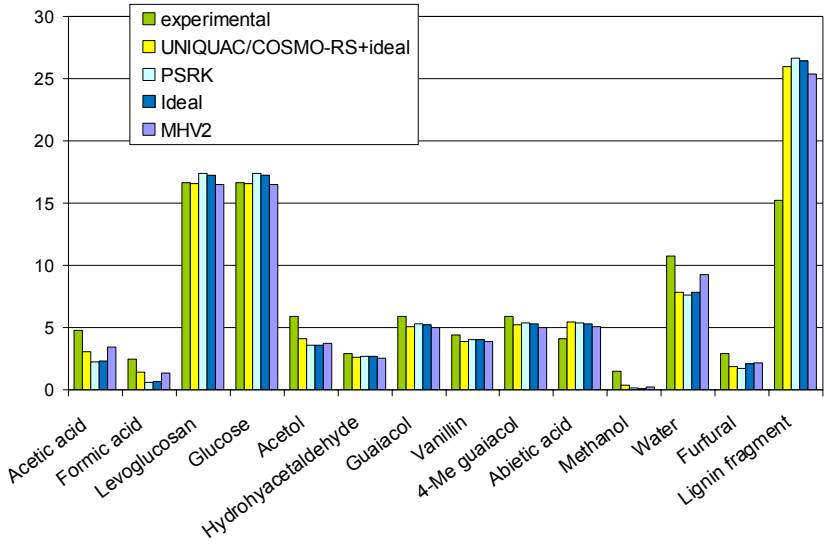

Figure 3.3. Weight percents of the compounds in product steam from the first condenser at 47°C.

Even though recalculated weight percentage of the compounds diverges only slightly from reconstructed experimental composition of the pyrolysis product, accuracy of the modelling remains moderate. The weight percent results are smoothed by high amount of heavy component (lignin fragment) and overall possible changes of small component contents are very low. Deviations of the model predictions from the experiments are more noticeable for low boiling compounds when weight fractions of each component condensed at some unit are plotted (Figure 3.4). One can see that in accordance with the calculations levoglucosan, glucose, abietic acid and lignin fragments are totally condensed at the first condenser whereas in the experiments it is not so due to extended formation of aerosols. These Figures (3.3 & 3.4) present only results for the first condenser. When concentrations are calculated for two condenser units together, where elecectric precipitator is modelled by following flash calculation at room temperature, the modelled and the experimental composition are almost the same because all the compounds are fully condensed after the two stages and no information about the condenser selectivity can be obtained.

Condensation at 4 different temperatures of sieve tray column was performed within Linfors experiments: at 47, 60, 76 and 87°C. With the selected models we simulated the 1 stage condensation at reported temperatures followed by flash condensation at room temperature. Resulting comparisons of predicted by the modelling water content in the product pyrolysis oil and organic loss versus the first condenser temperature in the process are shown in Figures 3.5 & 3.6. Composition of the feed at all experiments slightly varies and it has little effect on water content in the product and noticeable effect on the organic loss in the condenser.

For comparison purposes we also added to the figure the data from another pyrolysis experiments made by Westerhof *et al.* (2007). In the Westerhof experiments concentration of water in the reactor feed was higher and different condition of the reactor was used that can change significantly the product composition. Additionally two spray condenser where used sequentially for product collection after reactor. One can see on Figure 3.5 that modelled water content of the product do not depend on condenser temperature as it is for Lindfors data and for the Westerhof data. Organic loss in the Westerhof experiments increase with the condenser temperature, though for both Lindfors experiments and in the flash modelling, organic loss do not depend on the condenser temperature, because the second condenser operated at room temperature screens out temperature effect of the first condenser. Additionally the difference between the results of Westerhof and Lindfors could be expected due to different product collection principle, i.e. in the Westerhof experiment the product collection was based on liquid – liquid separation principle whereas in the Lindfors experiments it is based on vapour liquid equilibrium. However modelling of Lindfors experiments is not satisfactory, most likely due to made simplification, reconstruction of product and feed compositions and ignorance of aerosol formation.

One may conclude that the validation of the simple thermodynamic model (VLE, LLE approach) can not be made with the present pyrolysis oil experimental data due to complexity of the equipment and only approximate content known for both the condenser feed and the product.

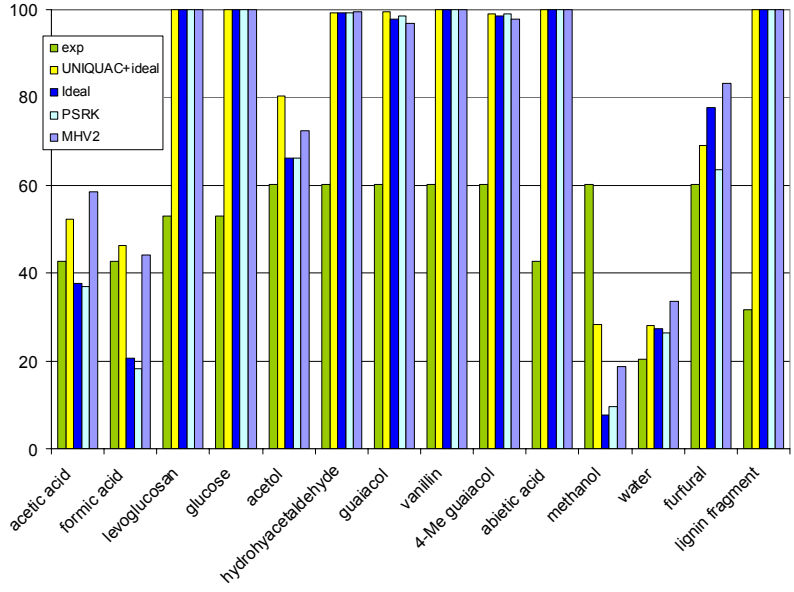

Figure 3.4. Weight percent of component recovery at first condenser stage at 47°C.

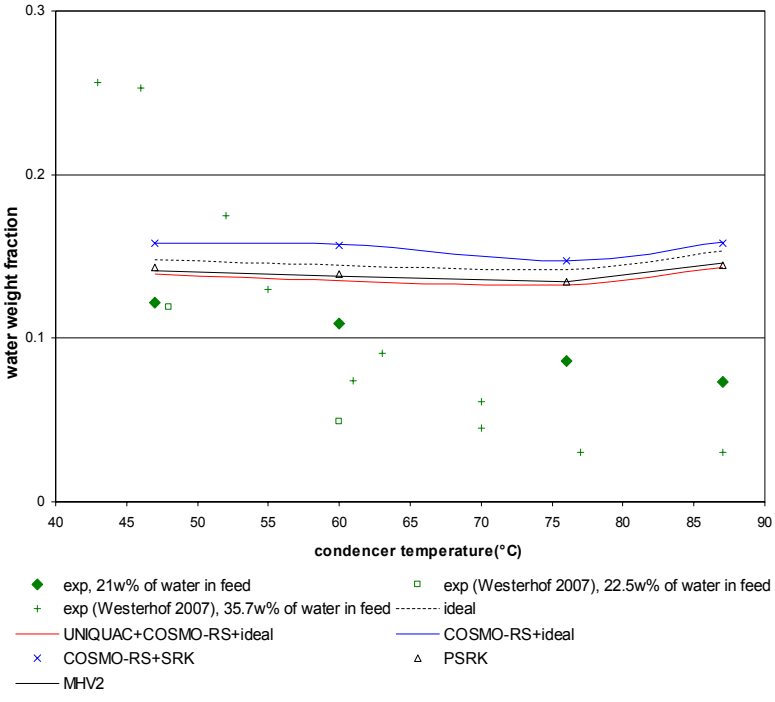

Figure 3.5. Water content in the product sum of two condensers.

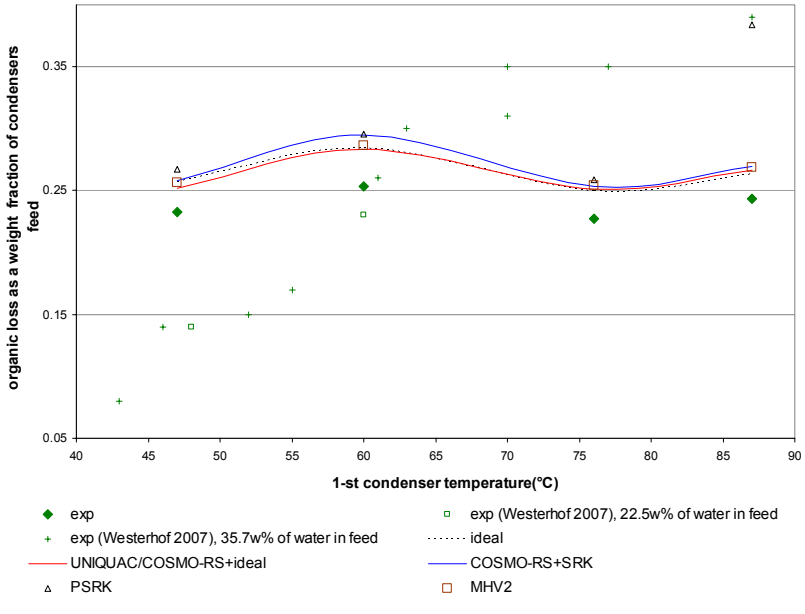

Figure 3.6. Organic loss after 2 condensation stages.

3.5.2 Ternary data

At previous section we described selection of models for the pyrolysis oil condensation (i.e. VLE and LLE equilibrium description) and we concluded that the concurrrent pyrolysis process experimental data can not be used for the model verification/validation. Therefore extensive search of ternary data was made for further testing of the models. Eight ternary aquatic systems were found. Summary of the found data is given in Table 3.5. One of the most difficult systems for modelling is a system where strong acid, water and aromatic compounds are combined, for example formic acid – water – furfural system. In Figure 3.7 one can see experimental and modelled data for this ternary system, binary and ternary data together. Combination of UNIQUAC model for liquid phase and Hayden O'Connell EOS for vapour phase was found to be the best. Accuracy of the ternary VLE data prediction with other models (PSRK, MHV2, UNIQUAC+ ideal vapour) was not satisfactory. It is worth noting that UNIQUAC coefficient were fitted using corresponding vapour description, i.e. Hayden O'Connell EOS. Generalization of the UNIQUAC + Hayden O'Connell description to all pyrolysis system however could be not so accurate. It is known that equations of state are not very accurate in description of heavy molecules behaviour, but there is no data available for validation of the model performance on heavy compounds from the pyrolysis list. In Figure 3.8 VLE diagram for acetic acid + water +methanol system predicted with MHV2 model is shown. The model does not use any fitted parameters. Utilization of UNIQUAC model with ideal vapour assumption or with Hayden O'Connell EOS gives just slightly more accurate prediction for those well known compounds.

Table 3.5. Ternary system experimental data for the selected pyrolysis oil compounds.

	2-nd component	3-d component	Data type	N sets	N points	Tmin, K	Tmax, K	Pmin, Mpa	Pmax MPa	References
1	acetic acid	CO_2	VLE	3	66	313.15	353.15	5	16	Bamberger *et al.,* 2004
2	acetic acid	furfural	LLE	3	15	298.15	308.15	0.101	0.101	Heric & Rutledge, 1960a & 1960b
	acetic acid	furfural	LLE	1	6	N/A	N/A	0.101	0.101	Pegoraro & Guglielmi, 1955
3	acetic acid	methanol	VLE	1	28	351.15	377.15	0.101	0.101	Sawistowski & Pilavakis, 1982
4	acetic acid	formic acid	VLE	2	56	375.15	382.15	0.101	0.101	Aristovich *et al.,* 1960; Conti *et al.,* 1960
	acetic acid	formic acid	VLE	1	16	309.15	315.55	0.006666	0.006666	Kushner *et al.,* 1966
	acetic acid	formic acid	VLE	1	16	341.15	345.15	0.026664	0.026664	Kushner *et al.,* 1966
	acetic acid	formic acid	VLE	1	49	283.15	381.15	0.101058	0.101058	Murayama, 1961
	acetic acid	formic acid	VLE	1	67	374.15	384.15	0.101	0.101	Wisniak & Tamir, 1977
5	CO_2	furfural	VLLE	1	3	303	343	5	5	Sako *et al.,* 1995
	CO_2	Methanol	VLLE	1	20	311	313	7	12	Yoon, *et al.,* 1993
6	formic acid	furfural	LLE	2	10	298.15	308.15	0.101	0.101	Langford & Heric, 1972
	formic acid	furfural	VLE	1	126	370.65	416.15	0.101	0.101	Tsirlin, 1969
7	furfural	methanol	VLE	2	78	-	-	0.101	0.101	Andreev & Tsirlin, 1954
8	Glucose	methanol	SLE	2	20	313.15	333.15	0.101	0.101	Peres & Macedo, 1997

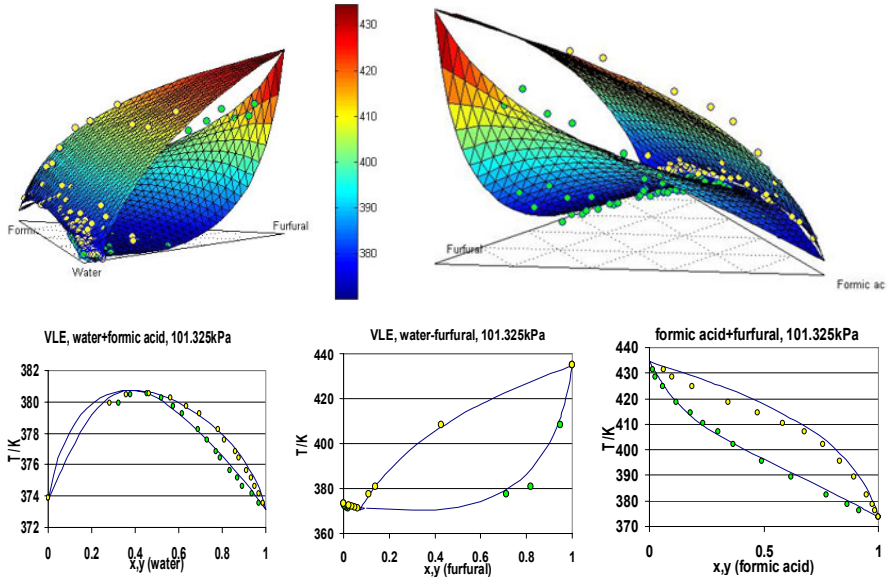

Figure 3.7. VLE diagram for formic acid – water – furfural system at 101.325 kPa; dots are experimental points (Tsirlin, 1969), lines are predictions, where UNIQUAC model is used for liquid phase description and Hayden O'Connell EOS is used for vapor phase.

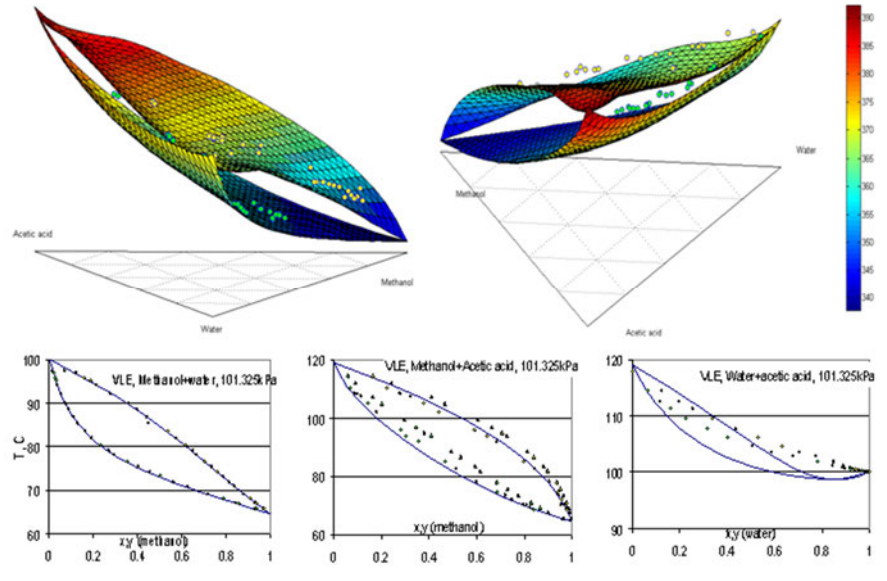

Figure 3.8. VLE diagram for acetic acid – water – methanol system at 101.325 kPa; dots are experimental points (Sawistowski & Pilavakis, 1982), lines are predictions, where MHV2 model is used for both phases description.

Further, the derived UNIQUAC model to predict liquid – liquid equilibrium was tested. Commonly parameters used for VLE can not be a applied for LLE description (for example separate models UNIFAC – VLE and UNIFAC – L LE) and no LLE data were used for fitting the UNIQUAC model parameters in our case (Table 3.4). Thus it was found that with developed set of the parameters no reliable liquid – liquid split can be predicted. For example LLE equilibrium in ternary system of furfural, water and formic acid (see Figure 3.9), where only at low overall water concentration the model predicts LL split, but composition of the phases deviate considerably from the experimentally determined. In contrast UNIFAC – LLE model predict relatively accurate phase compositions at low overall water concentration. However the UNIFAC – LLE model can not be used for the condenser liquid splitting prediction because it does not have group parameters for gaseous compounds.

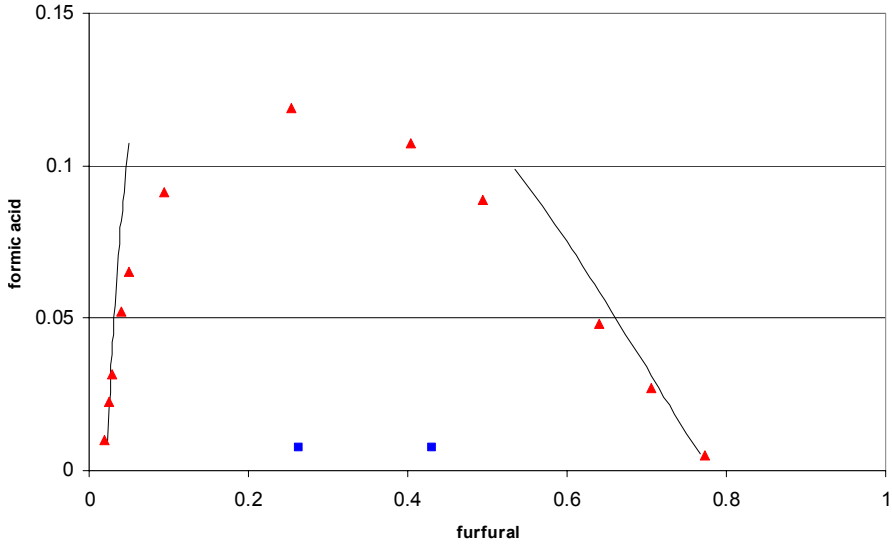

Figure 3.9. LLE equilibria of furfural, water and formic acid mixture. Red triangles are experimental data (Langford & Heric, 1972), blue squares are phase compositions predicted by UNIQUAC model with parameters given in Table 3.4, black solid lines are UNIFAC LLE model predictions.

3.6 Conclusions

Several models were tested to describe pyrolysis oil condensation. Predictive models were used due to the shortage of measured properties for pyrolysis compounds. The models were COSMO-RS and Predictive Soave Redlich Kwang (PSRK) equation of state, Modified Huron Vidal EOS (MHV2). UNIQUAC model was used combining COSMO-RS predicted data with several experimentally available binary system data. An equation of state (SRK, Hayden O'Connell) and the ideal vapour assumption were used for vapour phase description.

The modelling of the pyrolysis oil condensation shows that there is concurrently not enough experimental data available for development of a quantitative simulation of the condensation process or for verification of the parameters. Development of a flexible thermodynamic model which would be able to describe simultaneously polar, associating and polymeric compounds as well as gaseous compounds will further require more extensive multi-component equilibrium measurements, which may not be obtainable from industrial test runs that incorporate also other process phenomena (i.e. aerosol formation, mass transfer, reactions etc.).

4. Pyrolysis oil hydrodeoxygenation process

4.1 Hydrodeoxygenation process

The easiest way of applying pyrolysis oil as a transportation fuel is to combine it with diesel fuel. The properties that affects negatively on the usage of pyrolysis oil as transportation fuel is low heating value, solid content, high viscosity, incomplete volatility, corrosiveness and thermal instability (Czernik & Bridgwater, 2004). Because of the limited applications of pyrolysis oil an upgrading to a higher quality fuel is necessary in many cases. This is accomplished by removal of oxygen to the amount that the oil has sufficient properties for its combustion and/or re-forming the components in it (Bridgwater & Cottam, 1992). There are many ways of upgrading pyrolysis oil. The focus of this work was on the hydrodeoxygenation (HDO) of pyrolysis oil. HDO is a process where oxygen containing compounds react with hydrogen forming water and carbon dioxide in presence of a catalyst. The process occurs at elevated temperature and pressure. The elevated pressure is needed to keep the water in liquid form. The high pressure assists the reaction by making it faster through the bigger amount of dissolved hydrogen in the pyrolysis oil.

Thus thermodynamic models that can describe gas solubility at high pressure are needed for the process description together with mass transfer and kinetic modelling. The case study also serves to illustrate the methodology to be used while developing process models for chemically reactive unit operations in biorefining.

4.2 VLE and hydrogen solubility modeling

An equation of state (EoS) that can model the behavior of the process component interactions is essential for a good final result. Accordingly, three EoS, which are SRK, PSRK and PC-SAFT have been selected to model the thermodynamics of the hydrodeoxygenation process of pyrolysis oil. The EoS at hand was chosen for their dissimilarity, as SRK is a basic analytical cubic EoS, PSRK is an analytical cubic EoS with an activity coefficient model calculated through the Gibbs excess energy and PC-SAFT is a molecular based EoS.

4.2.1 Soave-Redlich-Kwong EoS

Soave modification of Redlich-Kwong (SRK) EoS (Soave, 1972) is a widely used model that has numerous further modifications. In this work two modifications were used, of which the first will be referred to as SRK. The SRK model differs from the original Soave modification by having an improved liquid molar volume correction for the EoS by Peneloux *et al.* (1982) and the water properties (enthalpy, entropy, Gibbs energy and molar volume) through the newer "NBS Steam Tables" (Haar *et al.*, 1984). Standard mixing rules are used. The model is presented through Equations (4.2.1–4.2.10).

$$p = \frac{RT}{V_m - b} - \frac{a}{V_m(V_m + b)} \tag{4.2.1}$$

$$a = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j \sqrt{a_i a_j}(1 - k_{ij}) \tag{4.2.2}$$

where

$$k_{ij} = k_{ji}$$

$$b = \sum_i x_i b_i \tag{4.2.3}$$

$$a_i = 0.42748\frac{R^2 T_{c,i}^2}{p_{c,i}}\alpha_i \tag{4.2.4}$$

$$b_i = 0.08664\frac{RT_{c,i}}{p_{c,i}} \tag{4.2.5}$$

$$\alpha_i = \left[1 + m_i\left(1 - T_{r,i}^{0.5}\right)\right]^2 \tag{4.2.6}$$

$$m_i = 0.480 + 1.574\omega_i - 0.176\omega_i^2 \tag{4.2.7}$$

$$V = V_m - c \tag{4.2.8}$$

$$c = \sum_i x_i c_i \tag{4.2.9}$$

$$c_i = 0.40768\left(\frac{RT_{c,i}}{p_{c,i}}\right)\left(0.29441 - z_{RA,i}\right) \tag{4.2.10}$$

where $z_{RA,i}$ is the Rackett liquid molar volume parameter.

Required parameters for the SRK are T_c, p_c and ω, which are presented in Table 5.2.1. For non-ideal mixtures also the binary interaction parameters (BIPs) should also be present. List of used binary interaction parameters is given in Appendix A. Accuracy of the parameters fitting and experimental data used for the fitting are also given in the Appendix A.

4.2.2 Predictive-SRK

PSRK model is the second extension of the SRK model. The mixing rule for parameter *a* is of Holderbaum and Gmehling (1991), which is a modification of the Michelsen-Huron-Vidal first order (MHV1) (Michelsen, 1990) mixing rule, shown in Equation (4.2.11). A modified UNIFAC activity coefficient model (Gmehling *et al.*, 1997; Horstmann *et al.*, 2000; Horstmann *et al.*, 2005) is used to represent the molar excess Gibbs free energy G_Y^E. For parameter *b* the linear mixing rule is used. The alpha function is the one of Mathias-Copeman, which is an extension of the alpha function in the original SRK is presented in Equation (4.2.12).

$$\frac{a}{bRT} = \sum_i x_i \frac{a_i}{b_i RT} - 1.546\left(\frac{G_Y^E}{RT} + \sum_i x_i \ln\frac{b}{b_i}\right) \tag{4.2.11}$$

$$\alpha_i = \left[1 + c_{1,i}\left(1 - \sqrt{T_{r,i}}\right) + c_{2,i}\left(1 - \sqrt{T_{r,i}}\right)^2 + c_{3,i}\left(1 - \sqrt{T_{r,i}}\right)^3\right]^2 \tag{4.2.12}$$

where $z_{RA,i}$ is the Rackett liquid molar volume parameter.

Required parameters for the SRK are T_c, p_c and ω, which are presented in Table 4.1. For non-ideal mixtures also the binary interaction parameters (BIPs) k_{ij} should also be present. List of used binary interaction parameters is given in Appendix A. Accuracy of the parameters fitting and experimental data used for the fitting are also given in the Appendix A.

4.2.3 Perturbed-chain statistical associating fluid theory

PC-SAFT equation of state developed by Sadowski and co-workers (Gross & Sadowski, 2001; Gross & Sadowski, 2002a; Gross & Sadowski, 2002b; Gross *et al.*, 2003). The main idea of the model is through reduced Helmholz energies to divide the intra-molecular forces into attractive and repulsive contributions, which is shown in Equation (4.2.13).

$$a^{res} = a^{hs} + a^{chain} + a^{disp} + a^{assoc} + a^{polar} \tag{4.2.13}$$

The repulsive interactions are described by a reference term and attractive interactions are described by a perturbation term.

The two first terms on the right hand side of Equation (13) describe the hard-chain fluid ($a^{hc} = a^{hs} + a^{chain}$), which accounts for the reference system in the model, that consist of chain molecules as a series of freely-jointed spheres. The reference fluid contribution has no attractive interactions and is based on the first order perturbation theory of Wertheim (Wertheim, 1983a; Wertheim, 1983b; Wertheim, 1985a; Wertheim, 1985b) that Chapman *et al.* (1990) applied to account for mixtures. For the reference system two pure component parameters are needed, which are segment number (m) and segment diameter (σ).

The dispersion term (a^{disp}) uses the perturbation theory of Barker and Henderson (Barker & Henderson, 1967a; Barker & Henderson, 1967b), which has been modified to account for chain molecules. The parameters needed for the non-associating and non-polar components are the segment number (m), segment diameter (σ) and the depth of the pair potential (ε/k), which are presented in Table 4.1. For mixtures the Berthelot-Lorentz mixing rules are applied in conjunction with binary interaction parameters (k_{ij}), which are shown in Equations (4.2.14) and (4.2.15).

$$\sigma_{ij} = \frac{1}{2}(\sigma_i + \sigma_j) \tag{4.2.14}$$

$$\varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j}(1 - k_{ij}) \tag{4.2.15}$$

The association (a^{assoc}) term is responsible for the short-range interactions. For these interactions two additional pure component (i) parameters are needed, which are the association energy ε^{AiBi}/k and effective association volume ($K^{A_iB_i}$) between association sites A_i and B_i. The mixing rules are for the association interactions those of Wolbach and Sandler (1998), shown in Equations (4.2.16) and (4.2.17).

$$\varepsilon^{A_iB_j} = \frac{1}{2}\left(\varepsilon^{A_iB_i} + \varepsilon^{A_jB_j}\right) \tag{4.2.16}$$

$$\kappa^{A_iB_j} = \sqrt{\kappa^{A_iB_i}\kappa^{A_jB_j}}\left(\frac{\sqrt{\sigma_{ii}\sigma_{jj}}}{(1/2)(\sigma_{ii} + \sigma_{jj})}\right)^3 \tag{4.2.17}$$

The equations of the dipolar (a^{polar}) term are developed by Jog and Chapman (1999).

Extension of the PC-SAFT model for copolymers, i.e. different types of homopolymers bound in the chain, is done through introducing into the dispersion term three interaction parameters of which two are homopolymer – solvent interactions and are not dependent of the copolymer solution. The latter two parameters were predetermined in AspenPlus software. The third parameter is dependent of the two monomer segments and is introduced equivalently as the binary interaction

parameter in Equation (4.2.15). The parameters were optimized for each binary system in Aspen Plus (AspenTech, 2010). The parameters are given in Appendix A.

Table 4.1. Parameters of the selected HDO process compounds used in SRK, PC-SAFT models.

	Tc / K	Pc / Mpa	omega	Mathias – Copeman parameters			PC-SAFT parameters		
				c1	c2	c3	m	σ	ε/k
Hydrogen	33.2	1.3	-0.22	0.1285	0.2477	-1.483	0.573	4.219	28.158
Methane	190.6	4.6	0.008	0.4926	0	0	0.965	3.747	152.943
Carbon dioxide	304.2	7.38	0.239	0.8255	0.1676	-1.704	2.647	2.563	149.138
Water	647.3	22.05	0.344	1.078	-0.5832	0.5462	3.044	2.031	309.881
Methanol	512.6	8.1	0.559	1.43	-0.6656	-0.1258	5.169	2.173	196.733
Ethanol	514	6.14	0.635	1.491	-0.402	0	6.022	2.288	185.553
1,2-Ethanediol	720	8.2	0.5254	1.258	0	0	5.069	2.481	274.556
Acetic acid	594.4	5.77	0.4218	1.296	-1.052	0.9223	3.961	2.782	254.451
Formic acid	588	5.81	0.3222	1.085	-0.8888	0.8129	2.559	3.038	314.077
Methyl acetate	506.8	5.0	0.3664	0.8556	-0.153	0.6005	3.15	3.222	234.643
Ethyl acetate	523.2	3.83	0.363	1.041	-0.1769	0.4951	3.324	3.413	237.224
Cyclohexane	553.8	4.08	0.213	0.8408	-0.3985	0.9415	2.51	3.86	279.367
Cyclohexene	560.4	4.91	0.2741	0.9815	-0.844	1.11	2.477	3.808	285.042
Methylcyclo-hexane	572.1	3.48	0.2477	0.93	-0.822	1.822	2.661	4.023	281.621
Benzene	562.1	4.89	0.212	0.8356	-0.375	0.9715	2.502	3.638	284.634
Toluene	591.7	4.11	0.257	0.9469	-0.5896	1.213	2.859	3.719	283.042
Benzaldehyde	695	4.66	0.32	1.288	-2.956	6.387	3.28	3.562	313.316
Catechol	770	6.50	0.5145	1.243	0	0	6.965	2.553	254.912
Guaiacol	697	4.73	0.2103	0.8608	-0.5773	1.4013	4.601	3.221	278.05
Phenol	694.2	6.13	0.44	1.524	-3.47	7.5	4.149	3.045	284.523
o-Cresol	697.6	5.0	0.443	1.288	-1.463	3.251	4.255	3.196	282.268
m-Cresol	705.8	4.56	0.464	1.165	-0.286	1.332	4.446	3.194	282.517
p-Cresol	704.6	5.15	0.515	1.54	-3.04	7.206	4.476	3.143	282.253

4.2.4 Model evaluation

For validation of the considered models extensive search of available experimental binary and ternary data was made. Total number of collected binary mixtures was 142 of 253 possible binary combinations of selected compounds (Table 2.1). The goal was to obtain data for the mixtures at conditions similar to those of the pyrolysis oil HDO process. Only VLE binary data, either bubble point data (xTp) or complete Tpxy data, are useful for HDO process modeling. That limitation is in many cases the reason for the lack of the data in question.

It was decided to classify the binary mixture data into groups depending on the component functional groups. Seven groups have been identified for the purpose of this work, namely light gases (hydrogen, methane and carbon dioxide), acids (acetic and formic acid), phenols (phenol, catechol, guaiacol and cresols), water, alcohols (methanol, ethanol and 1,2-ethanediol), esters (methyl and ethyl acetate) and hydrocarbons (cyclohexane, cyclohexene, benzaldehyde, benzene and toluene). When binaries are divided into groups (Table 4.2) the data is partly overlapping, i.e. hexane – hydrogen is included in two groups. In Table 4.2 the percentage differences of the residual root mean square errors (RMSE) between the different models is presented. PC-SAFT and SRK EOS are clearly better models, with some deviations in accuracy for the different compound groups. Thus for the best description of VLE several EOS (i.e PC-SAFT & SRK) have to be used depending on the compound classes.

Table 4.2. The percentage differences of the residual root mean square error between PC-SAFT, PSRK and SRK models.

Group	% of binary systems for which the 1-st model accuracy is higher		
	$RMSE_{PC\text{-}SAFT}<RMSE_{PSRK}$	$RMSE_{PC\text{-}SAFT}<RMSE_{SRK}$	$RMSE_{SRK}<RMSE_{PSRK}$
Light gases	84%	44%	82%
Acids	67%	61%	78%
Phenols	69%	40%	63%
Water	81%	69%	56%
Alcohols	60%	47%	64%
Esters	63%	38%	75%
Hydrocarbons	74%	61%	61%

4.2.5 Concentration of hydrogen

Though all mentioned above models can satisfactory predict VLE for gaseous compounds, higher precision of the gas solubility prediction is needed for the main reactant of HDO process, i.e for hydrogen. Hexadecane is considered as a solvent for calculation of the hydrogen concentration. Additionally it is assumed that the hydrogen concentration in the bulk hexadecane solution is constant and homogeneous, firstly due to the large hydrogen pressure and secondly due to the powerful mixing of the solution.

The hydrogen pressure for the experiments were in all cases 50 000 Pa. Neglecting nonideality of the vapour phase, hydrogen concentration in liquid phase can be calculated in accordance with Equation (4.2.18). Where Henry's coefficients ($H_{2,1}$) of hydrogen in hexadecane at HDO process temperature range are taken from work of Cukor and Prausnitz (1972).

$$H_{2,1} x_2 = P \tag{4.2.18}$$

4.3 Kinetics of the hydrodeoxygenation process

The kinetics is the next step in the investigation of the pyrolysis oil hydrodeoxygenation (HDO). The aim of the ongoing work is complete kinetic study of the HDO process providing all sufficient data for the kinetic modelling, i.e description of the reactive compound diffusion to the catalyst surface and the surface reactions as well as validation of the modelling against available experimental data.

4.3.1 Reaction pathway

The reaction of guaiacol is assumed to follow the reactions routes presented in Figure 4.1. The proposed reactive compounds and the reaction pathways are based on analysis of HDO experiments made by Heikki Harju (Harju, 2010). Source model compound for the HDO process is guaiacol. Its reactions can proceed by two different pathways, namely a primary route and a secondary route.

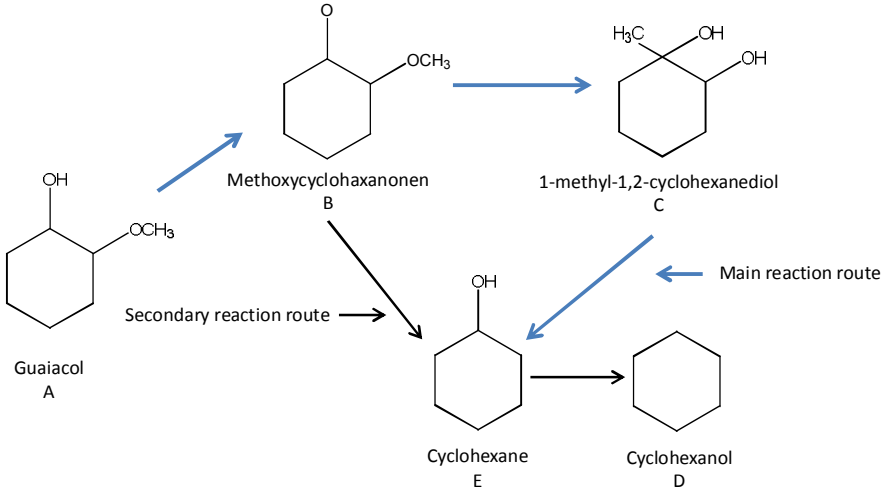

Figure 4.1. The assumed reaction pathways of guaiacol in the present experiments.

The reactions included in the primary reaction route are accounted in Equations (4.3.1–4.3.4). The reactants and reaction products are illustrated as letters A to E in order of the main reaction route, i.e. A stands for guaiacol, B stands for methoxycyclohexanone etc.

$$A + 2H_2 \rightleftharpoons B \quad (4.3.1)$$

$$B + H_2 \rightleftharpoons C \quad (4.3.2)$$

$$C + 2H_2 \rightleftharpoons D + H_2O + CH_4 \quad (4.3.3)$$

$$D + H_2 \rightleftharpoons E + H_2O \tag{4.3.4}$$

The calculations for the kinetic model were assumed to include the internal diffusion and reaction. Because of the good stirring and mixing of the reaction fluid the external diffusion was neglected.

Two types of reactions rates were tested for the calculations, namely a simple power law reaction rates that depends on the reactant, its reaction order and the reaction rate coefficient. The other reaction rate expression is a more complicated Langmuir-Hinshelwood reaction rate based on mechanism of Horiuti and Polanyi (1934).

4.4 Mass transfer and reaction rate

The kinetic model rate equation can be seen in Equation (4.4.1).

$$\frac{\partial c_i}{\partial t} = \frac{D_{e,i}}{\varepsilon_p r_p^2}\left[\frac{\partial^2 c_i}{\partial \lambda^2} + \frac{B+1}{\lambda}\frac{\partial c_i}{\partial \lambda}\right] + r_i \frac{\rho_p}{\varepsilon_p} \tag{4.4.1}$$

The kinetic model theorem is retrieved from the work of Lylykangas (Lylykangas, 2004). The dynamic model equation consists of two parts: diffusion rate is a first term of the right side of eq. 1 and the reaction rate is a second term of the eq. 1. In the equation lambda (λ) is a relative to centre position in the catalyst particle, calculated as ratio of the distance from the catalyst particle centre to the catalyst particle radius. The parameter B is the shape factor of the catalyst particle, which is described in Equation (4.4.2), were A_p and V_p are the catalyst surface area and catalyst volume, respectively. Due to the little knowledge of the catalyst, the parameter B was assigned the value of 3, which is a common value for crushed and sieved fractions of catalysts (Lylykangas, 2004).

$$B = \frac{A_p}{V_p} r_p \tag{4.4.2}$$

4.4.1 Power law reaction rate equation

The first tested model was a power law reaction that depends on the reactant, its reaction order and the reaction rate coefficient. The basic form of the expression is displayed in Equation (4.4.3). In accordance with Equation (4.4.4), i.e. linearized Equation of 1, the rate coefficient can be obtained from the point of interception of the logarithm of experimental reaction rate with the y-axis and the reaction order is a slope of line of the reaction rate logarithm.

$$-r_A = k' C_A^n = -\frac{dC_A}{dt} \tag{4.4.3}$$

$$ln(-r_A) = ln\left(-\frac{dC_A}{dt}\right) = ln(k' C_A^n) = ln(k') + n ln(C_A) \tag{4.4.4}$$

For computational reasons the values of the reaction order and the reaction rate coefficient was not always obtained by using all of the data points. This was due to the nonlinearity of the data points. The reaction rate coefficient and the reaction order were calculated by minimization of the Pearson product square that is presented in Equation 4.4.5. Value of the square of the Pearson product for one experimental set was usually over 0.9 and in few cases smaller.

$$r = \frac{\Sigma(x - \bar{x})(y - \bar{y})}{\sqrt{\Sigma(x - \bar{x})^2 \, \Sigma(y - \bar{y})^2}} \tag{4.4.5}$$

4.4.2 Langmuir-Hinshelwood type reaction rate equation

The second reaction model that was introduced in this work employs the Langmuir-Hinshelwood reaction rate and more specifically assumes the mechanism of Horiuti-Polanyi (Horiuti & Polanyi, 1934). This model is evidently much more complex than the one presented earlier.

In Appendix B the solution for the rate expression for reaction 4.3.1 is derived. Rate expressions for the reactions 4.3.2–4.3.4 are derived in the same manner, but final expression are much alike due to difference in the type of reactions.

In accordance with mechanism of Horiuti-Polanyi hydrogen is adsorbed without dissociation to the catalyst surface from where it reacts further with the reactant. The first addition of hydrogen to the reactant is assumed to be the rate limiting step.

Determination of correct values for the different rate coefficients (k_i) and equilibrium constants (K_i) is difficult optimization task due to complexity of the kinetic rate expressions (Appendix B). Those values can be obtained either by linear regression or by linearization as it was done in the section of power law reaction rate.

Finally at the end of Appendix B truncated forms of the rate expressions are presented for reactions 4.3.1.–4.3.4, where concentrations of intermediate products were removed from the expressions as they are not available in the experiments.

4.5 HDO process experimental data

The kinetic models calculations are based on data obtained from the Master's thesis work of Heikki Harju (Horiuti & Polanyi, 1934). In his experiments pyrolysis oil model compound guaiacol was hydrodeoxygenated (HDO) in batch reactor at 175°C, 200°C, and 220°C. Dilute solution of hexadecane was used as a solvent. The experiments were conducted for all three temperatures two times over a time period of 100 min and 120 min respectively. The catalyst used in the experiments

was rhodium over a zirconia oxide carrier. The experimental data used for the optimization the kinetic parameters are presented in Appendix C. For more information about the experiments the reader is referred to the previously mentioned work of Heikki Harju.

4.6 Results and discussion

In the kinetic modelling the HDO of guaiacol data was fitted against the power law reaction rate model (Equations (4.4.3–4.4.4)) and against Langmuir-Hinshelwood type reaction rate model (Appendix B). The Langmuir-Hinshelwood type model shows good correspondence of reactive compound concentrations change with time (see Figure 4.2). The concentration profiles at different catalyst depths (i.e. curves with the different lambda values) change in accordance with bulk composition. But absolute value of the concentrations is shifted to a lower concentration level as a result of combined effect of the diffusion and the reactions.

There are two types of the concentration curves, namely the curves that start at zero bulk concentrations (for reactions 4.3.1–4.3.4) and curves with finite initial bulk concentration (i.e. reactions 4.3.1). The latter type is the more informative because it determines the diffusion rate compared to the reaction rate. In particular at the beginning of the experiment the reagents and the products concentrations inside the catalyst are zeros and the components diffusion rates can easily be determined for the reaction 4.3.1. For all reactions the concentration inside the catalyst peaks immediately and fast after the beginning of the experiments. Then the concentrations start to decrease. When the reactions are neglected the concentration curves for the different positions inside the catalyst particle are shifted just a little to a higher concentration level. This demonstrates that the overall rate is determined by the reaction rate and is not considerably influenced by the internal diffusion. In Figure 4.2 the difference between the diffusion and reaction rates for reaction 4.3.1 calculated with power low model is shown. The experimental data used for the calculations are given in Appendix C, Table 1. Other examples of the rate differences at different temperatures are provided in Appendix D.

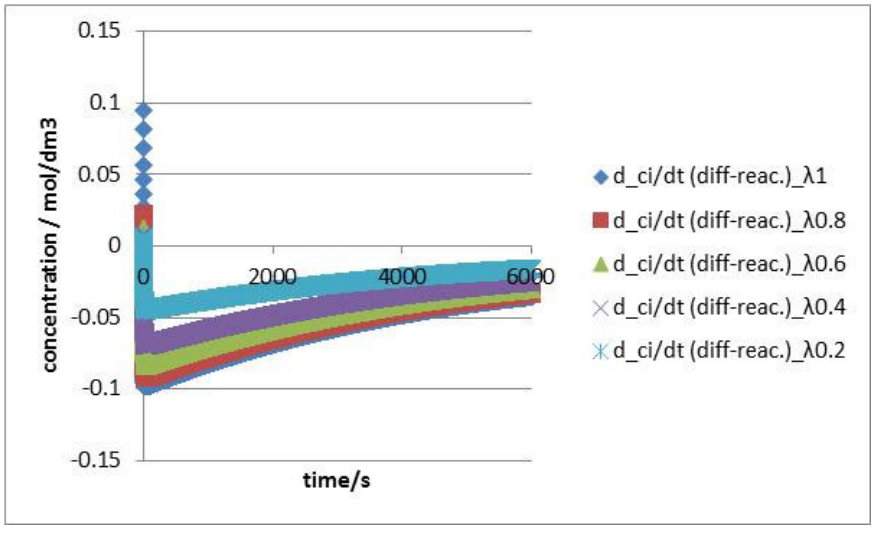

Figure 4.2. The difference between diffusion rate and power law reaction rate for the experiment described in Table 1 in Appendix B.

The reaction rate inside the catalyst depends on the reagent concentrations. The concentrations in turn depend on the diffusion rate. This is the reason why the reaction rate is slow at the beginning of experiment. After the concentrations reached their maximum for given lambda value, the reaction rates reach their own maximum and then slow down with moderate rates. From the figures in Appendix D one can see that the biggest deviations of the reagent concentrations can be observed at maximum or minimum of the bulk concentrations independent of the catalyst depth (the lambda value).

Reactions 4.3.1–4.3.4 start from zero reagent concentrations inside the catalyst and accelerate at the beginning of the experiments. It was found that after the reaction rates reach maxima only the reactions 4.3.2 notably decrease. Correspondingly that reaction does not progress significantly after first hour. 1-methyl-1,2-cyclohexanediol is the main product after first hour and its concentration remains more or less constant. This indicates that the chosen reaction conditions are not favourable for the reaction to proceed fully toward cyclohexane.

The dependence of the reagent concentrations at most inner catalyst layers (lambda = 0.2) on bulk concentration is weak. This is due to the assumption that in the centre of the catalyst the concentration is always zero and also due to low catalyst discretization, i.e. division of the catalyst particle to few thick layers (big $\Delta\lambda$).

From the shown figures one can see that the primary pathway is dominant. Concentration of methyl-1,2-cyclohexanediol is much bigger compared to any of its product concentrations and the methyl-1,2-cyclohexanediol is the only compound that does not participate in the reaction of secondary pathway.

5. Thermal hydrolysis for levulinic acid production

Hot water and steam are typically used for processing wood and biomass to gain intermediate biorefining liquors from the lignocellulosic raw material. Hydrolysis with hot water in mild condition extracts hemicelluloses from the wood. Steam explosion and harsh conditions will break hemicellulose and cellulose chains to the fermentable sugars. Then, the thermal hydrolysis is somewhere between those treatments. The aim of thermal hydrolysis is to hydrolyse cellulose, but create larger molecules than sugars. With properly controlled kinetics, it is possible to decompose the cellulosic polysaccharide molecule to 4-oxopentanoic acid and formic acid monomers. The 4-oxopentanoic acid, also called levulinic acid is then a possible product of thermal hydrolysis with optimal conditions. As side products, furfural and formic acid are formed and the residue of the lignocellulosic biomass will remain as solid bio-char.

Levulinic acid is seen as a platform chemical with several possible uses. The most obvious end-use will be as fuel additives. Options with more added-value are fragrances, plasticizers and resins. See Figure 5.1.

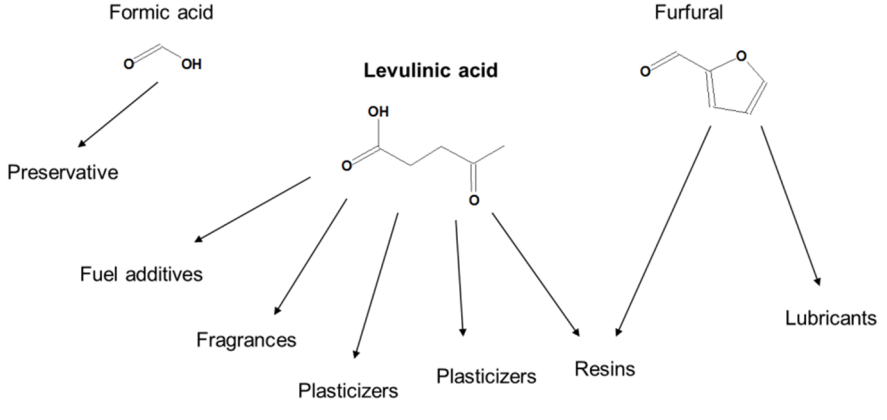

Figure 5.1. Main products of the process (Fitzpatrick, 1990; Fitzpatrick, 1997).

A commercial process called Biofine has been developed by Hayes *et al.* (2006). Proposed benefits for the process are: (i) Simple thermo-chemical process with two-stage acid-catalyzed reaction, (ii) No microbes or enzymes needed, (iii) Versatile raw materials, e.g. wood residues, fuel crops, agricultural crops, food wastes, waste and reclaimed paper, and algea & seawater crops, (iv) High yields, from C6-sugars 50% to levulinic acid and 20% to formic acid, from C5-sugars 50% to furfural, (v) Residue is burnable bio-char, and (vi) an energy efficient process has been claimed.

An overview of the process is shown in Figure 5.2. The first reaction at higher temperature, ~ 220°C is very fast, ~ 10 seconds. The second reaction takes longer, ~ 20 min, at little lower temperature, 190°C. Several separation steps are requires before levulinic acid, furfural, formic acid, bio-char and several other minor components are fractionated.

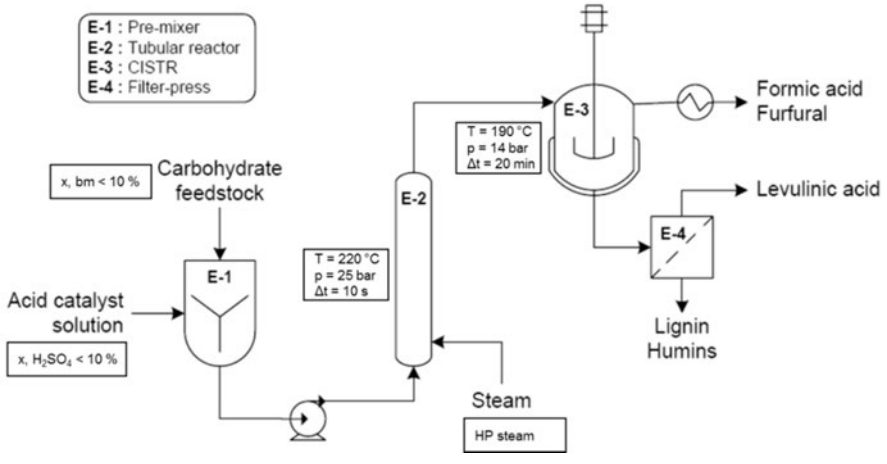

Figure 5.2. Proposed production process (Fitzpatrick; 1990; Fitzpatrick, 1997; Hayes *et al.*, 2006; Girisuta, 2007).

The main interest in the BIOSCEN project was to study the thermal-hydrolysis reaction and fractionation of end products with modelling tools. The feasibility of VLE (flash/distillation/evaporation) as a separation technique was studied. As the proposed production process is thermo-chemical with rather high temperatures and pressures, a thermodynamic equilibrium approach was chosen and later extended with as a constrained equilibrium model. ChemSheet was used as the modelling software. In addition, ICAS 12 was used for predicting thermodynamic properties of some substances.

5.1 Simplified model with pure substances

First model is based on the properties of pure substances and ideal phases, which are shown in Table 5.1. Water, levulinic acid, formic acid, furfural and nitrogen are in the gaseous phase. Water, levulinic acid and formic acid are in the aqueous phase. Liquid phase is assumed for furfural. Glucose and xylose are solid phases which are here representing the biomass.

Results are shown in the Figure 5.3. Based on this model, the evaporation temperatures of furfural, formic acid and levulinic acid are different and it could be possible to use VLE to separate formic acid and furfural from levulinic acid. This is also proposed by the original authors of Biorefine process (Fitzpatrick, 1990).

Table 5.1. Species and phases used for the simplified model. Ideal mixtures assumed. Part of the thermodynamic data is from literature, part of it is estimated by ICAS 12.

		Href[298]	Sref[298]	Cp
Gas	H2O(g)	Factsage	Factsage	Factsage
(IDMX)	N2(g)	Factsage	Factsage	Factsage
	Formic_acid(g)	DIPPR801	DIPPR801	ICAS12
	Levulinic_acid(g)	DIPPR801	DIPPR801	ICAS12
	Furfural(g)	DIPPR801	DIPPR801	ICAS12
Aqueous	H2O(a)	Factsage	Factsage	Factsage
(IDMX)	Formic_acid(a)	DIPPR801	DIPPR801	ICAS12
	Levulinic_acid(a)	DIPPR801	DIPPR801	ICAS12
Pure phases	Furfural(l)	DIPPR801	DIPPR801	ICAS12
	Glucose(s)	DIPPR801	DIPPR801	(constant)
	Xylose(s)	DIPPR801	DIPPR801	(constant)
Constrains	Rx			
	Measured value	Estimated value		

5

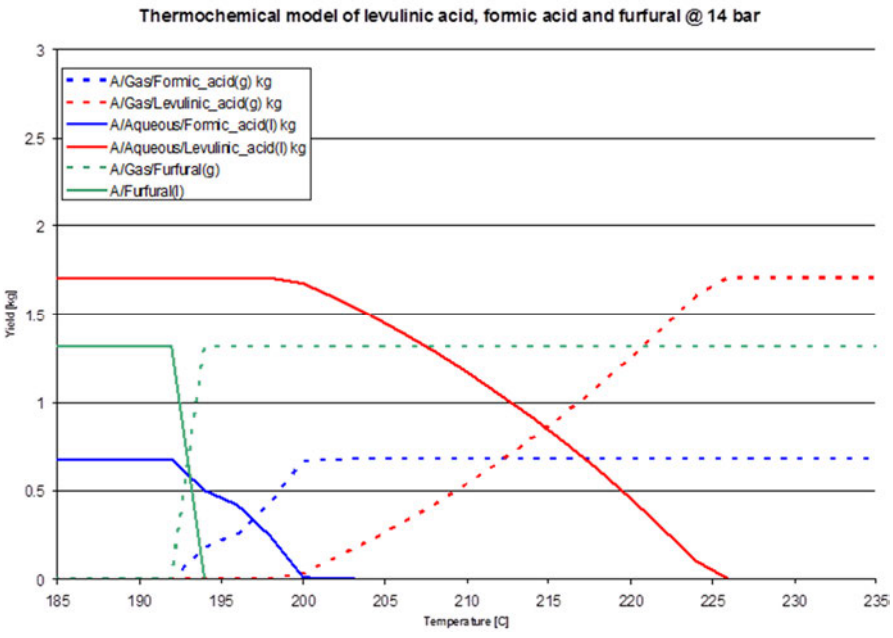

Figure 5.3. A schematic thermochemical model of the acid-furfural system using ideal mixtures.

5.2 Kinetic models of thermal hydrolysis

As earlier mentioned, the thermal hydrolysis process is controlled by the temperatures, sulphuric acid addition and reaction times in two reactors. Several reaction kinetics models are presented in the literature (Girisuta 2007; Chang *et al.*, 2006; Chang *et al.*, 2009). An example of the yield of levulinic acid as function of temperature and H_2SO_4-wood ratio is shown in Figure 5.4.

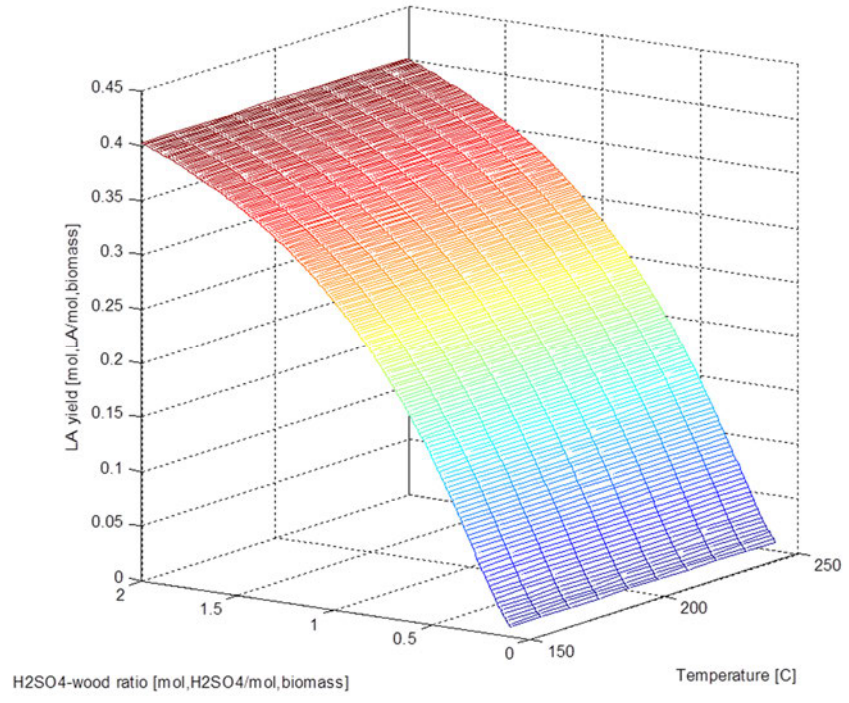

Figure 5.4. Reaction kinetics of thermal hydrolysis. literature (Girisuta, 2007; Chang *et al.*, 2006; Chang *et al.*, 2009).

5.3 Refined model with non-ideal mixtures

Once a simplified model showed promising results, the thermochemical model was refined to include unideal mi xtures and interactions with different species. First a Unifac binary-model for formic acid and levulinic acid was evaluated. The system is presented in Table 5.2.

Table 5.2. Binary system of levulinic acid and water.

Phase	Species	C	H	O
Gas	Levulinic acid	5	8	3
Gas	Water	0	2	1
Liquid	Levulinis acid	5	8	3
Liquid	Water	0	2	1

The results of this binary system were validated against literature data. Results agreed well in low 10 torr pressure, see Figure 5.5. Predicted values at 1 and 15 bar are in Figure 5.6.

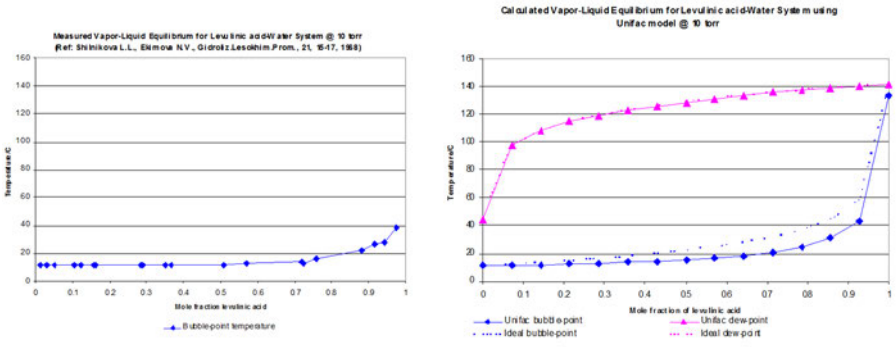

Figure 5.5. Bubble and dew point of levulinic acid – water system at 10 torr. Reference data on left and predicted values on right.

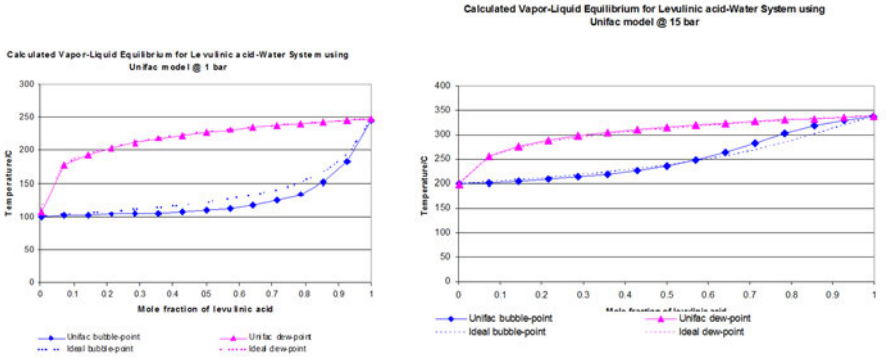

Figure 5.6. Bubble and dew point of levulinic acid – water system at 1 and 15 bar. Predicted values.

Next model described the formic acid – water binary system. Unifac and virial models were applied here as well. The system is presented in Table 5.3.

Table 5.3. Binary system of formic acid and water.

Phase	Species	C	H	O	Formic*
Gas	Formic acid	1	2	2	1
Gas	$(HCOOH)^2$	2	4	4	2
Gas	Water	0	2	1	0
Liquid	Formic acid	1	2	2	1
Liquid	Water	0	2	1	0

The results of this binary system were validated against literature data. Results agreed rather well in 1 bar pressure, see Figure 5.7. Predicted values at 15 bar are respectively shown in Figure 5.8.

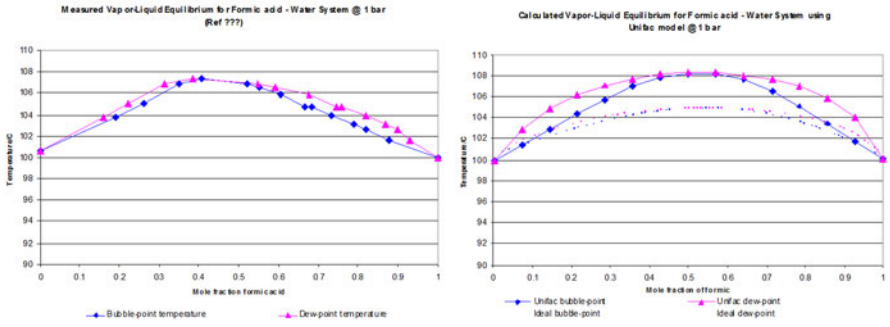

Figure 5.7. Bubble and dew point of formic acid – water system at 1 bar. Reference data on the left and predicted values on the right.

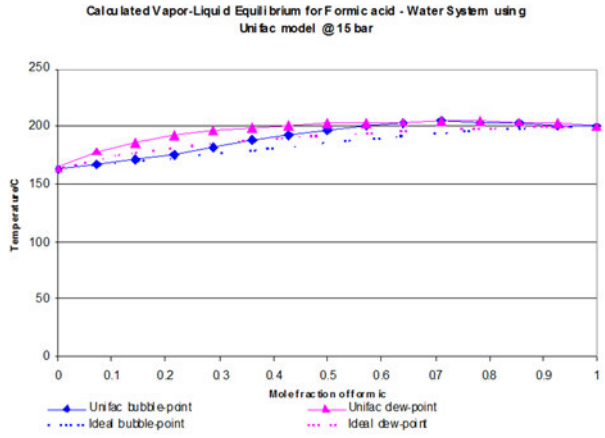

Figure 5.8. Bubble and dew point of formic acid – water system at 15 bar. Predicted values.

Two binary systems were incorporated into a ternary system of levulinic acid – formic acid – water. System is shown in the Table 5.4. Predicted bubble and dew points of this ternary system are shown in Figure 5.9. No reference data for this system was found.

Table 5.4. Ternary system of levulinic acid, formic acid and water.

Phase	Species	C	H	O	Formic*
Gas	Formic acid	1	2	2	1
Gas	$(HCOOH)^2$	2	4	4	2
Gas	Levulinic acid	5	8	3	0
Gas	Water	0	2	1	0
Liquid	Formic acid	1	2	2	1
Liquid	Levulinic acid	5	8	3	0
Liquid	Water	0	2	1	0

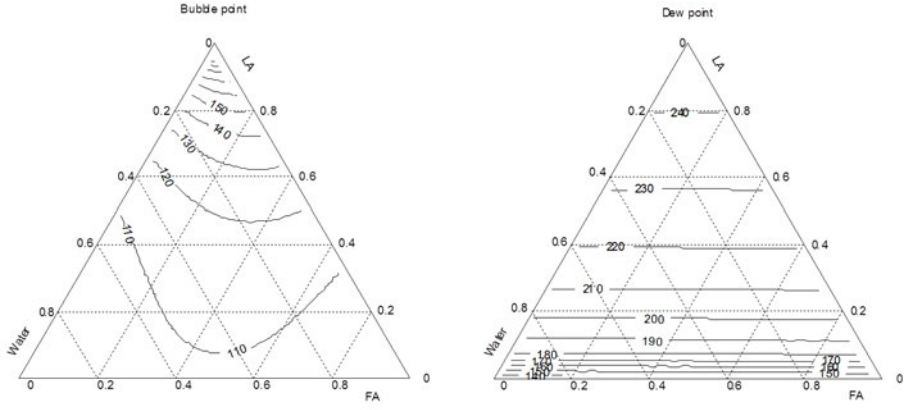

Figure 5.9. Ternary plots of levulinic acid – formic acid – water system at 1 bar. Bubble on the left and dew point on the right. Values are predicted.

Finally, the kinetically restricted decomposition of glucose was included in the model. Chemical system is shown in Table 5.5. Constraint Glucose_R as well as additional system component Rg was included in Table 5.5. By controlling the constraint it is possible to define how much glucose is decomposed. A prediction is presented in Figure 5.10 with reference data from the literature.

Table 5.5. Decomposition of glucose included.

Phase	Species	C	H	O	Formic*
Gas	Formic acid	1	2	2	0
Gas	$(HCOOH)^2$	2	4	4	0
Gas	Levulinic acid	5	8	3	0
Gas	Water	0	2	1	0
Liquid	Formic acid	1	2	2	0
Liquid	Levulinic acid	5	8	3	0
Liquid	Water	0	2	1	0
Solid	Glucose_R	6	12	6	0
Solid	Glucose	6	12	6	1

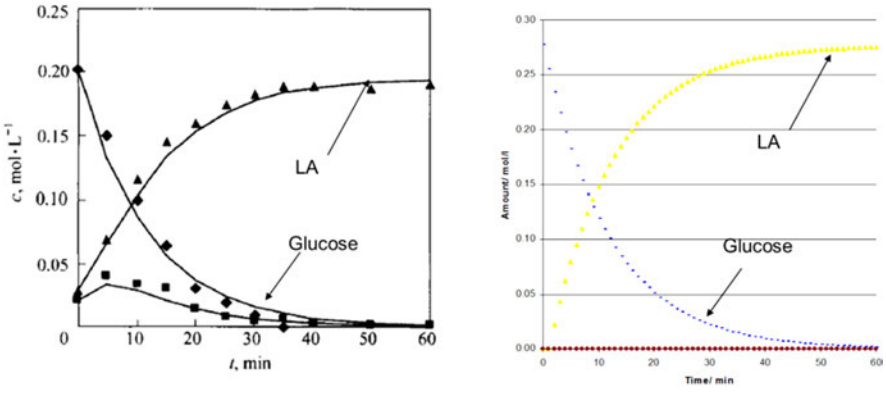

Figure 5.10. Kinetically constrained decomposition of glucose. Reference (Chang *et al.*, 2006) on the left and predicted values on the right.

5.4 Conclusions

In this study the vapour-liquid equilibrium of levulinic acid-formic acid-(furfural)-water system was studied. First pure substances and ideal mixtures were applied. Model was refined, and non-ideal mixtures utilized Virial and Unifac parameters. Later on kinetic constraint was added in order to model the decomposition of glucose.

The thermal hydrolysis and levulinic acid itself is increasingly popular research subject. However there seem to be very little experimental data available even for the binary systems. Predictive modelling technique such as Unifac group contribution method and different methods utilized by ICAS 12 software can fill some gaps in knowledge. The final answer of the feasibility of this proposed process is difficult to deduce without further experimentation. It is yet evident that extractive methods instead of mere fractionation distillation should be adapted.

6. Dilute acid pretreatment

Cellulose hydrolysis to monosaccharides is a major preceding step for the production of biofuels (ethanol, buthanol) and related chemicals via the fermentation route from lignocellulosic raw material. Then, before the actual hydrolysis most often catalysed by enzymes, a pretreatment stage is required. A mild treatment with e.g. dilute acid is generally preferred to provide suitable raw material for the biochemical hydrolysis.

An overview of the mechanisms of dilute acid pretreatment of lignocelluloscis is given by Palmquist and Hahn-Hägerdal (2000), or Baugh and McCarty (1988). Jacobsen and Wyman (2000), provide a review of models existing in literature about dilute acid pretreatment, divided by models for hemicellulose and glucose hydrolysis.

Models for hemicellulose hydrolysis are generally governed by first order reactions and Arrhenius equations for the adjustment of kinetic parameters to temperature and acid concentration. The main differences between models found in literature lie in the level of detail. The simplest model considers reactions from hemicellulose directly to xylose and further to degradation processes. The second level of detail would be to distinguish between fast and slow hydrolysing hemicellulose (Esteghlalian *et al.,* 1996). Third and most detailed option is to include oligomers as a reaction step between hemicellulose and xylose (Garrote *et al.,* 2001; Jacobsen & Wyman, 2000). Degradation products from xylose are usually unelaborated.

For glucose the consensus in literature is again on a first order reaction from cellulose to glucose. Amorphous cellulose, which is instantly hydrolyzed to glucose in dilute acid pretreatment, is usually included as an initial glucose concentration (Bhandari *et al.*, 1984; Saeman, 1945). The refinement here is again the step through olisaccharides (Jacobsen & Wyman, 2000). Glucose degradation products are more specifically investigated to be hydroxymethylfurfural (5-HMF), which is further degraded to levulinic acid and a negligible amount of formic acid (McKibbins *et al.,* 1962; Palmqvist & Hahn-Hägerdal, 2000).

For this work the kinetic model equations and parameters for hemicellulose hydrolysis for the matlab file were taken from Esteghlalian *et al.* (1996), since it gave a thorough analysis of poplar wood as well as corn stover in addition to all kinetic parameter values necessary, including a review on previously found kinetic parameters in literature for comparison. Cellulose hydrolysis however is not covered

in this paper and was therefore taken from Bhandari *et al.*, 1984, who also considers corn stover as substrate. A model for the degradation of glucose is provided by McKibbins *et al.* (1962).

The overall model structure compiled from these papers and implemented in matlab is as shown in Figure 6.1. All reactions are modelled by first order kinetics. The kinetic parameters are adjusted to temperature and acid concentration through Arrhenius equations and are raw material specific.

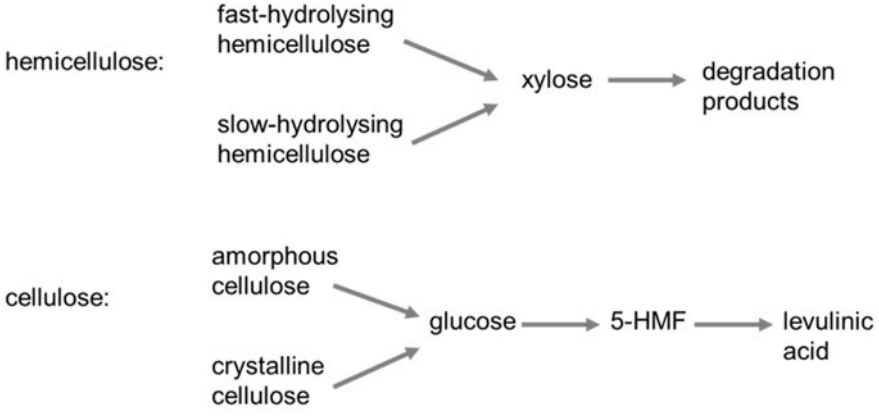

Figure 6.1. Overall model structure for the dilute acid pretreatment.

The model is valid for a temperature range from 140 to 180°C and an acid concentration between 0.4 to 1.5% (Bhandari *et al.,* 1984; Esteghlalian *et al.,* 1996; McKibbins *et al.,* 1962). Depending on the raw material, the xylose is hydrolysed after 20–40 minutes reaction time. After this time, cellulose will continue to slowly hydrolyse and degradation products will be formed.

Specifications for temperature, nominal acid concentration, solid ratio of the loaded mass and the neutralization ability of the substrate have to be provided to the model. The kinetic parameters in the implemented model are for hydrolysis of corn stover, because no complete set of parameters for wood material was available. Output of the model are the concentrations of xylan, xylose, xylose degradation products and further cellulose, glucose, 5-hmf and levulinic acid over time. Figure 6.2 shows an output plot of an example simulation over the first 20 minutes of the reaction.

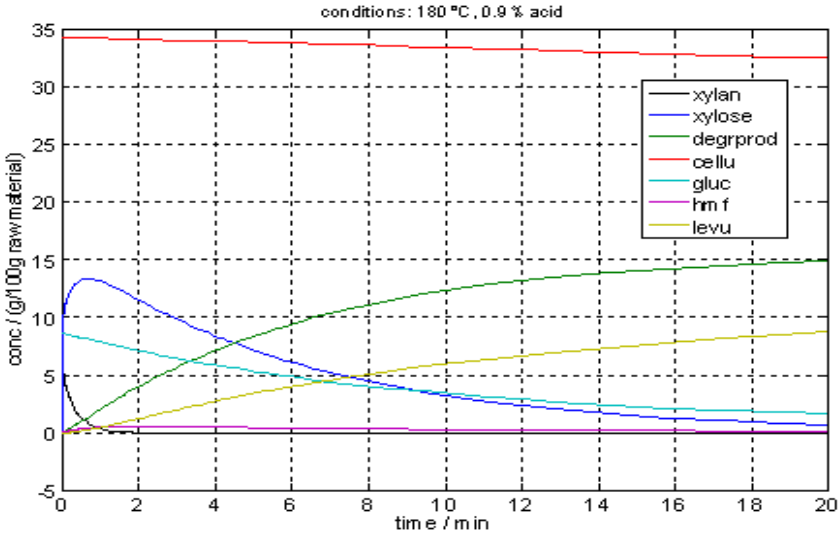

Figure 6.2. Example of simulation results for dilute acid pretreatment of corn stover.

7. Enzymatic hydrolysis of lignocellulosic biomass

Enzymatic hydrolysis is an essential step in the saccharification of lignocellulosic biomass. Its applications range from the partial hydrolysis of cellulose to the total hydrolysis of cellulose and hemicelluloses. The former yields a fibrous product that contains sugar monomers, while the latter yields a carbon source for the production of a carbon based chemical in a biorefinery.

A process such as enzymatic hydrolysis is very much dependent on water as a reactant, but the presence of excess water makes the recovery of product energy intensive. High-consistency hydrolysis has been proposed to alleviate these problems and ultimately make the process economical in large scale (Hodge *et al.,* 2009). Thus a salient parameter to be considered in the hydrolysis process will be the ratio of solids and water.

The mathematical model of the enzymatic hydrolysis of dilute-acid pretreated lignocellulosic biomass first presented by Kadam *et al.* (2004) and later modified by Zheng *et al.* (2009) is one of the most readily available dynamic models in this area. The former model contained parameters for the inhibition by xylose, while the latter contained parameters for the apparent inhibition caused by lignin.

In this work, the two models were merged. The merged model accounted for enzyme loading, process time, and inhibition by glucose, cellobiose, and xylose. However, the merged model still seemed inadequate with respect to temperature, pH, and solid loading. Furthermore, the mechanisms of the enzyme classes are known to far greater detail (Beckham *et al.,* 2011; Igarashi *et al.,* 2011) than any of the current mathematical descriptions portray.

With adequate enzyme dosage, most substrates can achieve nearly total hydrolysis at a solid loading of 10 g/L (Zhang *et al.,* 2007). As the solid loading is increased, the relative yield is reduced (Kristensen *et al.,* 2009). The net effect may be due to a number of reasons, but if it is understood, the process may be improved.

The modeling objective was to create a mathematical description capable of reproducing the trends and approximate magnitudes of significant phenomena under select process conditions and expected feed variability. The model performance is commensurate with all models referenced by Zhang and Lynd (2004) in terms of mechanistic detail and process parameters.

7.1 Exclusions

A number of phenomena were excluded from the modeling effort. At low solid loading, the yield increases with increasing mixing rate (Wang *et al.,* 2011). Since the mixing is extremely different at high solid loading, the effect of mixing rate was excluded from the model. The model assumes adequate mixing. Thus, the free solvent phase was assumed yielding a uniform and homogeneous mixture, i.e. concentrations are defined and valid to describe the state of the system. Diffusion and convection were excluded.

Enzyme adsorption onto lignin is assumed to be reversible although it has been proven that irreversible adsorption does occur. The actual fate and influence of denatured enzyme is unknown. Therefore, the enzyme in the model is chosen to simply disappear when it denatures. Denaturation is assumed to be irreversible.

Surfactant adsorption to lignin is believed to prevent unproductive binding of enzymes to lignin, thereby producing higher yields and enabling better recycling of enzymes (Kristensen *et al.,* 2007). The use of polymers (Kristensen *et al.,* 2007; Börjesson *et al.,* 2007) or surfactants to selectively cover lignin patches Zhu *et al.,* 2009) was not included in the model.

7.2 Rheology

Water retention value is a standardized measure of the capability of a fibrous material to retain water within their structure (www.nordstand.com). When the water retention value is determined using negative pressure-assisted filtration, the result depends on the pressure used. The same phenomenon is partially responsible for the thixotropic and pseudoplastic effects observed in fiber suspensions. The mechanical pressure squeezes out water from the fiber, thereby lowering the viscosity of the suspension. The process takes finite time since the water molecules need to diffuse through the pores.

Furthermore, as the enzymes break down the fiber wall, water stored within the fiber becomes free. This is probably the main contribution to the apparent thixotropic effect observed during hydrolysis.

The effect of stirring rate (Wang *et al.,* 2011) has not been included. One possible mode of impact is a slight reduction in the effective water retention value due to increased dynamic pressure within the fiber. Shear forces may also affect denaturation (Reese & Ryu, 1980).

The model described herein relies heavily on thermodynamics of aqueous solutions, which requires that the concentration of a dissolved substance must be well defined. The concentration of a dissolved substance is/becomes defined when the dissolved particles of the same substance are statistically interchangeable. The particles are interchangeable if their solution shells are identical. In practice, this means that the water content of the suspension may not be too small.

7.3 Water

Water has multiple roles in enzymatic hydrolysis (Kristensen *et al.,* 2009). It is a reactant and it is needed for product desorption, both of which play a larger role at higher consistency. While a full thermodynamic model for the aqueous solution was implemented, a computationally faster approximation proved adequate as long as the pH remained relatively constant. The activity of water can be approximated from the amount of water, the amounts of solutes within the water, and the average hydration numbers of the solutes. Figure 7.1 shows an example of how the activity of water decreases with the extent of hydrolysis.

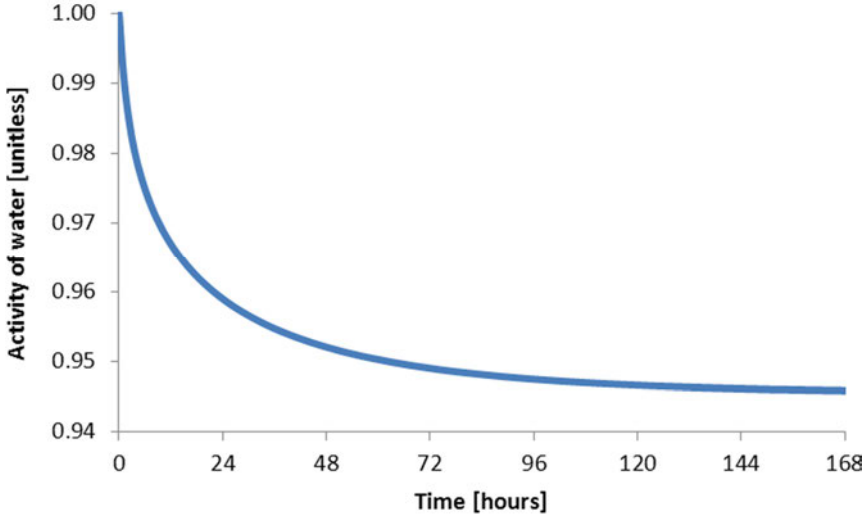

Figure 7.1. The activity of water decreases during hydrolysis.

Hemicelluloses and lignin naturally contain carboxylic acids, which have a negative charge when deprotonated. During pretreatment, some of these acids will detach from the fiber matrix. Depending on the conditions, alkaline or acidic, both lignin and hemicelluloses decompose and may form new acid moieties. Furthermore, adding an oxidizing agent containing oxygen will greatly accelerate the formation of carboxylic acids, especially during high-temperature pretreatment.

Since the slurry coming into the hydrolysis process is not yet liquefied, there will be some fiber character left. This fiber, even if it is only a loose polyelectrolyte spanned between cellulose bundles, contains a fixed charge. This charged fibrous structure will probably form a region of space where the solvent is different from the bulk solvent. This region is effectively modeled as an additional aqueous phase even if some of the region also can be classified as a surface. The fixed charge in this aqueous phase excludes ions with the same sign and attracts ions with the opposite sign, thus resulting in a different pH and ion distribution within

the fiber phase compared to the bulk solution phase. A negatively charged fiber phase has a higher pH than the external solution and it accumulates multivalent cations as described by Donnan partitioning (Koukkari *et al.,* 2002).

7.3.1 Water as a reagent

The mechanism of a hydrolytic reaction requires the participation and consumption of one water molecule. If the solvation shells of the participating molecules are included in the reaction stoichiometry, then the apparent water consumption may be larger. In this work, product desorption was modeled separately. Therefore, the hydrolytic reaction rates were modeled as a first order reaction with respect to water.

7.3.2 Water as a solvent

Engineering models typically scale abundances with the total mass of the system, e.g. in units g/(kg mixture). Since this ignores the effect of solvents, the models are rarely usable at other consistencies than that of the measurements. By making the abundances proportional to the solvent, i.e. using concentrations proportional to the amount of solvent (g/kg water), the liquid model becomes independent of the solid fraction and therefore scalable with solid loading. If the solid fraction remained constant throughout the hydrolysis process, this scaling would be scalar multiplication.

As referred above, cellulosic pulp tends to retain water and swell in the presence of water. If one assumes that the water within the fiber does not contain active enzyme, the loss of solvent can be modeled using an effective water retention value (WRV). As hydrolysis proceeds, the fibers along with their water retention capacity disintegrate. In the process of simultaneous saccharification and fermentation, the microbial cells also retain water even if the computational description of the amount of biomass is in cell dry weight (CDW). Any and all solutes affect the effective amount of water for solvation. Figure 7.2 shows the effect of WRV on the relative hydrolysis yield at higher consistencies.

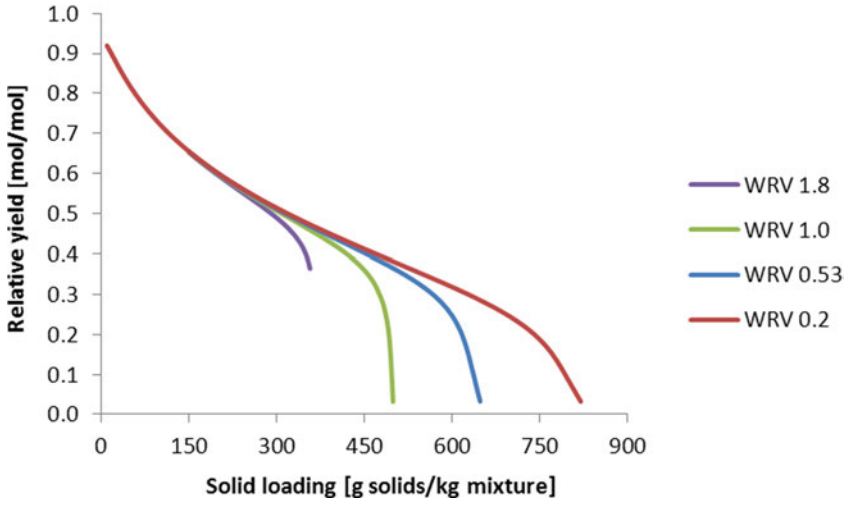

Figure 7.2. WRV influences the loss of relative yield at higher solid loading.

7.3.3 Approximating the activity of water

Let WRV_{pulp} quantify the relative amount of water inaccessible to enzyme due to the fibrous biomass. Let WRV_{cells} quantify the relative amount of water inaccessible to enzyme due to the microbial biomass. The free solvent amount f_W is then described by Equation (7.3.1), where t_W is the total amount of water in the mixture, m_{pulp} is the amount of dry fiber and m_{CDW} is the amount of dry cell material.

$$f_W = t_W - WRV_{pulp} * m_{pulp} - WRV_{cells} * m_{CDW} \quad (7.3.1)$$

The water retention value of the fiber depends strongly on the pretreatment method used. The water retention value of the micro-organism depends on the species and the general growth conditions, but has little impact on the overall result. The value for WRV_{cells} (7.3.2) was calculated based on the average dry mass content for yeast reported by Jørgensen *et al.* (2007).

$$WRV_{cells} = \frac{1-0.31}{0.31} \approx 2.23 \quad (7.3.2)$$

All solutes bind water molecules in their shell of solvation. The organic molecules here are far larger than most inorganic ions. Therefore, the traditional method of estimating the activity of solvents through mole fraction is not adequate. The bound amount of water was estimated based on average numbers of hydration. The activity of water a_W is then calculated by Equation (7.3.3), where n_k is the mole amount of solute k and H_k is the average hydration number for that solute.

$$a_W = \frac{1}{f_W}(f_W - \sum_k H_k n_k) \quad (7.3.3)$$

The average numbers of water molecules structured by a molecule of glucose, cellobiose, or xylose used to approximate the activity of water, were derived from Hollenberg and Hall (1983).

7.3.4 Approximating the activities of solutes

As explained in Chapter 7.3.2, the abundances are best described proportional to the solvent. While activity coefficients may vary quite a lot during the process, their effect may yet be assumed small in comparison to the uncertainty associated with kinetic parameters. Therefore, the activities of solutes were approximated by molal concentrations (7.3.4), typically in units mol/kg water.

$$c_{solute} = \frac{n_{solute}}{m_{solvent}} \tag{7.3.4}$$

7.3.5 Solid loading

Solid loading is the amount of solid lignocellulosic biomass introduced into the reactor. It may be as low as 1% in typical enzyme property studies, while it usually exceeds 30 mass-% in high-consistency hydrolysis process runs. By scaling the model according to the water content in the mixture rather than the total mixture mass, the model applicability range is improved.

All empirical process data included in the model were apparently measured at 8 w-% or 10 w-% consistency (Kadam *et al.,* 2004; Zheng *et al.,* 2009). Therefore, it was only a fortunate coincidence that some earlier versions of our model (Figure 7.2, low WRV) seemed capable of reproducing the relative yields for such a broad range (2 w-% to 40 w-%) of solid loadings as those measured by Jørgensen *et al.* (2007).

7.4 Overview of enzymatically catalyzed reactions

The enzymatic hydrolysis of cellulose is kinetically described by three four functionalities and by three four enzymes. The models by Kadam *et al. (*2004) and Zheng *et al.* (2009) considered three functionalities. The additional detail was required to model simultaneous saccharification and fermentation (SSF), where implicit treatment of endoglucanases were no longer an option. Parameter sensitivity analyses made for the first crude version of SSF clearly indicated that enzyme adsorption was far more important than in the earlier models for only enzymatic hydrolysis. This report contains the inferences of the more developed parts of these models.

Our models are made for the enzyme mixture Celluclast 1.5 and the additional enzyme Novozyme 188. Celluclast is a commercial preparation of cellulases produced by Trichoderma reesei. It contains at least two cellobiohydrolases and two endoglucanases. Novozyme 188 is a commercial preparation of β-glucosidase produced by *Aspergillus niger*.

T. reesei cellobiohydrolase I (Cel7A) is 65 kDa and cellobiohydrolase II (Cel6A) is 58 kDa [18]. *A. niger* β-glucosidase is 130 kDa (Sipos *et al.,* 2010). *T. reesei* endoglucanase I (Cel7B) is 55 kDa and endoglucanase II is 48 kDa (Suurnäkki *et al.,* 2000). The three major enzymes Cel7A, Cel6A, and Cel7B all have a cellulose binding domain (CBD), with which they attach to the surface of cellulose.

The progressive exoglucanase Cel7A split off units from the reducing end of glucan chains and often move along the chain to its next act of hydrolysis. The non-progressive exoglucanase Cel6A split off units from the non-reducing end of glucan chains. The endoglucanases Cel7B hydrolyze bonds in the backbone of glucan chains. All three enzymes catalyze the hydrolysis of β(1→4) linkages between D-glucose repeating units.

The molecular masses of reactants and products below indirectly describe the apparent reactions. The molar masses of water, glucose, and xylose were calculated based on their atomic formulas. The molar mass of cellobiose was calculated as the difference of two glucose molecules and a water molecule. The representative molar mass of a glucan repeating unit was calculated as the difference of a glucose molecule and a water molecule. The molar mass of a representative lignin repeating unit was defined as 180 g/mol. The representative molar mass of a xylan repeating unit was calculated as the difference of a xylose molecule and a water molecule.

7.5 Enzyme denaturation

It is well known that the so called optimal temperature of an enzyme is not really optimal under all conditions. The rate of reaction increases with temperature but so do thermal denaturation. Since the easily quantifiable portion of denaturation is irreversible, any fixed-length assay will yield a maximum in enzymatic activity. This apparent optimal temperature is a trade-off between increased activity of active enzymes and the faster accumulation of denatured enzyme.

The temperature optimum, *i.e.* the compromise between increased enzyme activity and increased denaturation, depends on the process time (Figure 7.3). The original model based on the Kadam *et al.* (2004) and Zheng *et al.* (2009) structure needed only the addition of enzyme denaturation and refitting the common apparent activation energy before it reproduced this trend. The fact that the modified model predicts a time-dependent temperature optimum is a significant improvement toward true temperature dependency.

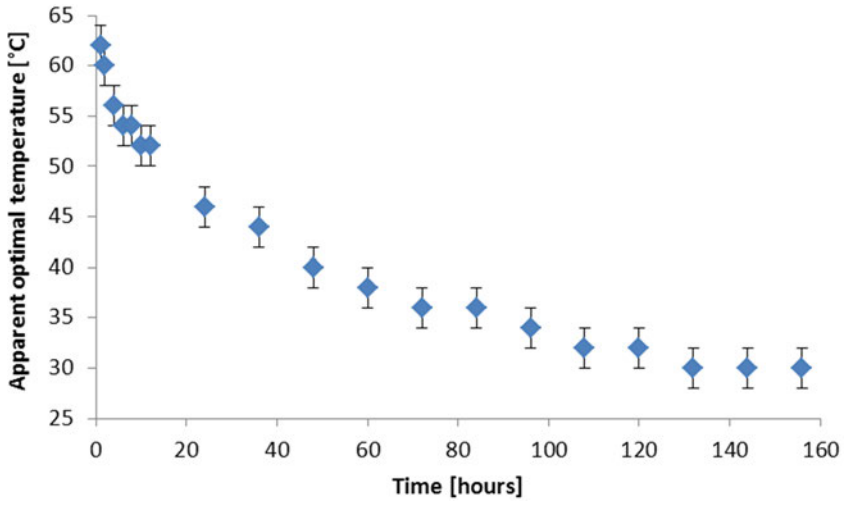

Figure 7.3. The optimal temperature of enzymatic activity depends on the length of the assay. The error bars represent the grid size of the calculations used. Assays are typically shorter than 2 hours, but this particular system required longer times due to other process parameters such as solid loading, pH, and total enzyme amounts.

Past modeling efforts have not achieved consensus on the mechanism of denaturation (Bansal *et al.,* 2009; Zhang *et al.,* 2010). Some assume a first order reaction, while others support a second order reaction (Zhang *et al.,* 2010). The issue is further complicated by the presence of multiple denaturation types. Thermal unfolding would be proportional to the amount of enzyme, thereby appearing as first order denaturation. The enzymes tend to have cysteines forming intermolecular sulfur bridges. Therefore, they are susceptible to intermolecular cross-linking, which appears as second order denaturation. Under suitable temperature and redox conditions, both may be reversible or effectively irreversible. Since neither temperature nor redox-potential is likely to change during the course of hydrolysis, any reversible denaturation would be impossible to separate from uncertainties in the amount and activity of the enzymes. Therefore, we chose to model only the effectively irreversible denaturation of enzymes using first order kinetics.

The most common data published in literature is the residual activity of enzymes that have spent a specific amount of time at a specific pH and temperature before readjusting to optimal pH and temperature for conducting the assay. Given the incubation time and the relative reduction in activity, the rate of denaturation can be calculated.

The pH and temperature dependent kinetic parameters were correlated to data published by Karboune *et al.* (2008) and Kwon *et al.* (1999). An example is shown in Figure 7.4. Due to the lack of precision and the relatively short exposure times of the measurements, the denaturation rates are relatively unreliable for long time

simulations. Particularly, the pH dependencies suffer from discontinuities in the data due to the different enhancing action of ions in different buffers.

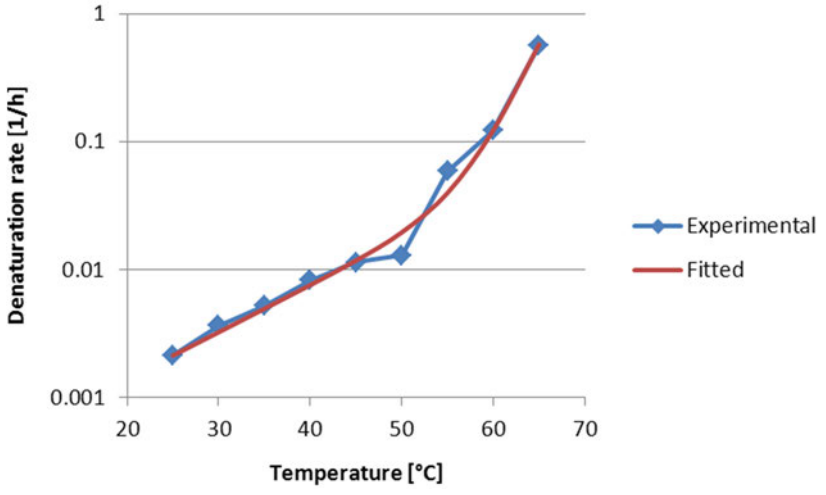

Figure 7.4. Temperature dependency of the denaturation rate for Cel7A. The data below 25°C was uncalculable due to noise.

The parameters for the progressive exoglucanase Cel7A was obtained by fitting to data derived from Figures 2C and 4C in Karboune *et al.* (2008) (*P. funiculosum* cellobiohydrolase activity, residual activity after incubation for 2 hours). The values for the non-progressive exogucanase Cel6A were assumed equal to the other exoglucanase. The parameters for the endoglucanase Cel7B were obtained by fitting to data derived from Figures 4C and 4D in Kwon *et al.* (1999) (*T. viride* endoglucanase I, residual activity after incubation for 30 min). The parameters for β-glucosidase were obtained by fitting to data derived from Figures 2B and 4B in Karboune *et al.* (2008) (*P. funiculosum* β-glucosidase, residual activity after incubation for 2 hours).

7.5.1 Temperature dependency

Due to the two-domain structure of these enzymes, the temperature dependency of denaturation was described by the Arrhenius-inspired correlation (7.5.1). There may be other explanations for the enhanced denaturation rate at high temperatures.

$$\rho_{\text{Temperature}} = e^{T_{\text{explow}}(T-T_{\text{reflow}})} + e^{T_{\text{exphigh}}(T-T_{\text{refhigh}})} \qquad (7.5.1)$$

where T is the temperature, T_{reflow} the reference temperature of domain 1, T_{refhigh} the reference temperature for domain 2, T_{explow} is a scale factor for domain 1,

$T_{exphigh}$ is a scale factor for domain 2, and $\rho_{Temperature}$ is the rate factor describing the temperature dependency.

7.5.2 pH dependency

Denaturation brought on by pH is essentially an acid- or base-catalyzed reaction. It is therefore expected to be exponential in pH and pOH (7.5.2). Low or high pH directly disrupts ion bridges and the neutralization of amino acid side chains mess with the electron density in and around the backbone. Since the electron density is important for protein folding, it is likely a dominant factor for protein stability as well. It can be well expected that the temperature and pH have a synergistic effect of aggravating each other. Since proteins contain both acids and bases, it can be assumed that there is a rough symmetry around some optimal pH. This is, however, not strictly true since the location and number of acids compared to bases is different.

$$\rho_{\mathrm{pH}} = \frac{e^{pH_{\mathrm{exphigh}}(pH-optpH)} + e^{pH_{\mathrm{explow}}(optpH-pH)}}{2} \tag{7.5.2}$$

where *pH* is the pH, *optpH* is an optimal pH, pH_{explow} is a scale factor for branch 1, $pH_{exphigh}$ is a scale factor for branch 2, and ρ_{pH} is the rate factor describing the pH dependency.

Determining reaction rates is rarely done without buffers and few buffers can be used for large pH ranges. The metal ions and the small anions will affect the denaturation rate much the same way they affect the catalysis rate if they fit into a void somewhere in, at, or near the active center. For every different buffer used, at least one offset parameter needs to be added during parameter regression. Therefore, a specific buffer should be used for at least two different pH values. Furthermore, buffer transitions should preferably be done such that there is at least one pH in common with the adjacent buffer series. This will fix the value of the absolute shift, which is the minimal added parameter. If there is a two-point overlap, the other parameter notifying the difference in scales can be quantified. In this work, we used only single parameter shifts wherever the shifts were obvious. The regressed parameters are therefore “valid” for the buffer ions in the buffer at optimal pH (for any pH value) and not for any other buffer ions in other buffers used (regardless of the original buffer ions used at any particular pH measurement).

7.6 Enzyme adsorption

The three enzymes Cel7A, Cel6A, and Cel7B have cellulose-binding domains (CBD) that help them attach to the cellulose surface (Suurnäkki *et al.,* 2000; Sipos *et al.,* 2010). Enzyme adsorption onto cellulose can generally be divided into reversible and irreversible adsorption (Ma *et al.,* 2008). Irreversible adsorption can

be captured by differential equations, but is hard to differentiate from enzyme denaturation. This section discusses fully reversible adsorption.

Adsorption of hydrolytic enzymes with or without binding domains onto cellulose can be modeled by the quantity of surface sites and the adsorption energy. If the adsorption energy is independent of surface coverage, then this depicts a Langmuir isotherm. The reversible adsorption can be assumed to be in equilibrium if the relaxation time constant is much smaller than the characteristic time constant for the subsequent hydrolysis reaction.

It is common to make a separate Langmuir adsorption isotherm for each enzyme-substrate pair. However, unless the enzymes adsorb to completely independent surface segments, they compete at least partially for available surface sites.

7.6.1 Lignin

Both endoglucanases and exoglucanases tend to condense onto any biomass surface. Bovine serum albumin (BSA) is able to reduce the amount of enzyme bound to substrate samples taken from different time points of hydrolysis (Heiss-Blanquet *et al.,* 2011). Cel7A lost 33% in the beginning, 28% thru most of the hydrolysis, and has only weakly bound enzyme in the end when there is 11% cellulose left. Apparently, it takes some time for Cel7A to dig in, but then it will not let go until the adsorption sites run out. They run out far before the amount of cellulose runs out indicating that the surface area of CBD+CD (cellulose binding domain and catalytic domain) is larger than the CBD alone. Cel7B behaves the same way as Cel7A; being constant at 25% reductions due to BSA until finally going up to 100% at the end with 11% cellulose. Cel6A seems to lose its lower energy adsorption spots much faster than Cel7A (Heiss-Blanquet *et al.,* 2011), which is consistent with the notion of Cel6A sharpening the fibers at the non-reducing end (Igarashi *et al.,* 2011).

While the hypothesis that bovine serum albumin, polyelectrolytes, and surfactants prevent nonproductive enzyme adsorption onto lignin may be true (Kristensen *et al.,* 2007; Börjesson *et al.,* 2007; Zhu *et al.,* 2009), it also has other often overlooked consequences. BSA and polyelectrolytes adsorb to the surface of cellulose and lower the surface concentration of cellulases. The consequences are two-fold. The lower surface concentration results in less congestion i.e. less traffic jams. For a progressive cellulase in motion, surfactants are probably easier to nudge aside than other hydrolytic enzymes. Thus, more of the surface is covered, but it is easier to traverse.

Palonen *et al.* (2004) measured zero enzyme adsorption onto lignin. Heiss-Blanquet *et al.* (2011) discovered that lignin from different sources have different affinity for enzymes. While Klason lignin embedded in the carbohydrate solid may be relatively inert (Palonen *et al.,* 2004), non-fibrous lignin seems to adsorb enzymes (Heiss-Blanquet *et al.,* 2011) and affect the overall rate of hydrolysis (Zheng *et al.,* 2009). Zheng *et al.* (2009) fitted the Langmuir adsorption parameters to experimental process data at the same time as the kinetic parameters.

7.6.2 Effective surface area

Cel7A and Cel6A adsorption onto lignocellulosic residues (Palonen *et al.,* 2004) clearly indicate that these enzymes adsorb onto cellulose and not lignin. The publication contains estimates of the lignin fraction of each substance and the distribution coefficients for each enzyme on each substrate. When the cellulose amount is overestimated as the non-analyzed fraction of the solids, the distribution coefficients line up perfectly on a specific curve. This curve has a fractal dimension so close to 0.5 that it is likely to have some physical reason for it. For an infinite cylinder, the surface area fraction is proportional to the radius while the volumetric fraction is proportional to the square of the radius. Thus, there is a physically reasonable explanation for that particular fractal dimension. If the dissociation constant is assumed constant, then the free surface area for adsorption of enzymes is proportional to the square root of the volumetric cellulose fraction (Figure 7.5).

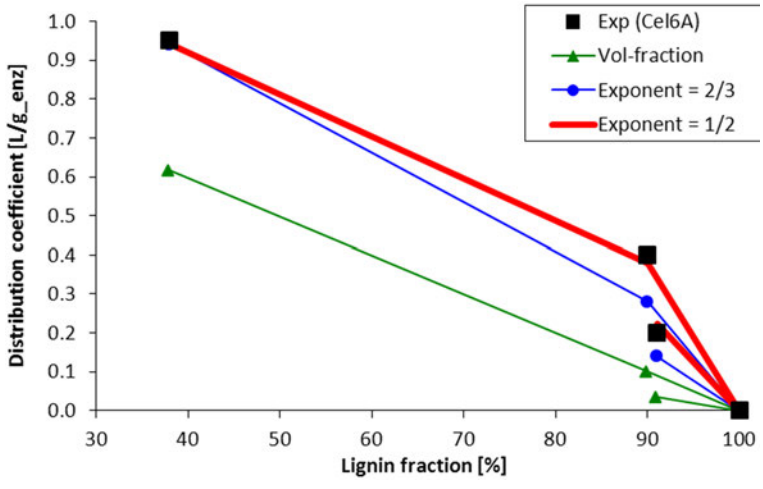

Figure 7.5. The distribution coefficient of Cel6A versus the fraction of lignin in the solid adsorbent. The exponent = 2/3 correspond to a spherical particle, while the exponent = 1/2 correspond to an infinitely long cylinder. The graph contains four data points in each of the four groups. The lines in between the points have no meaning except for aiding the visualization of the points in each group.

The protein content of the CEL-lignin having a data point at 92% lignin with lower distribution coefficient explains most of the deviance from the ideal curve (Figure 7.6). However, it does not explain all of it, so one could hypothesize that some hemicellulose or ash remains in that particular sample type. The discrepancy is but twice the size of uncertainty in the analysis, so the hemicellulose hypothesis is plausible. The particle size and shape may also be a contributing factor for Cel7A, which is indicated by the curve with exponent = 2/3.

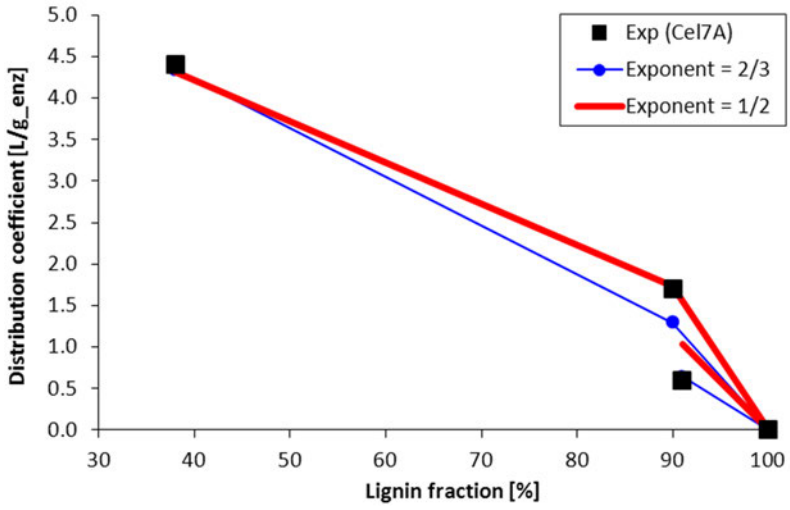

Figure 7.6. The distribution coefficient of Cel7A versus the fraction of lignin in the solid adsorbent. The exponent = 2/3 correspond to a spherical particle, while the exponent = 1/2 correspond to an infinitely long cylinder. The graph contains four data points in each of the three groups. The lines in between the points have no meaning except for aiding the visualization of the points in each group.

The results by Heiss-Blanquet *et al.* (2011) support a square root dependency between bound enzyme and cellulose fraction.

Considering a fairly dense mixture of fibers in water, not all fiber surfaces are equally well accessible. The available surface area is guesstimated to be roughly proportional to the volume fraction of water in the mixture.

7.6.3 Binding sites

Kim and Hong (2001) discovered that Cel7A and Cel6A each adsorb onto two different adsorption sites. The two-line Scatchard plots in (Kim & Hong, 2001) can only be reproduced by assuming that there are two mutually exclusive packing lattice structures. Dual Langmuir binding would not give two lines, but one smooth curve. The (Kim & Hong, 2001) data are for 30 min, while the data are for 120 min. It seems the surface needs a little more time to relax when two lattices are supposed to coexist on the surface.

The maximum binding capacity of each enzyme is roughly halved when they occupy the same surface (Kim & Hong, 2001). They simply share the surface. However, slightly denser packing is achieved, probably because the lattice blocks are slightly different in shape. Their respective catalytic domains may also adsorb to distinct surface sites the other can not use.

According to Table 2 in Palonen *et al.* (2004), the CBD will still find adsorption spots when the CD can not. The CBD is only 3–9 kDa compared to the approximately 50 kDa CD (Igarashi *et al.,* 2006; Suurnäkki *et al.,* 2000).

To sum up after Chapters 1.6.1 and 1.6.2, endoglucanases and exoglucanases adsorb onto the surface of cellulose. Cel7A adsorbs to the reducing end with both CBD and CD. Cel6A adsorbs to the non-reducing end with both CBD and CD. Cel7A, Cel6A, and Cel7B can adsorb anywhere in between using only their CBD. Cel7B can adsorb anywhere on an intact cellulose chain with both CBD and CD. If both domains are adsorbed, the enzyme is held more tightly on the cellulose surface. For numerical convenience, no enzyme is allowed to adsorb only by its catalytic domain (CD), although enzymes lacking a CBD certainly do just that.

Even if the enzyme amount in the solution has reached a near constant value, changes may still occur at the surface. The competitive binding isotherm in Kim and Hong (2001) is an indication that the surface has not had time to relax in 30 minutes. Figure 7.7 shows a hypothetical example of how the enzyme amount in solution may be nearly constant, but there are still significant changes occurring on the surface.

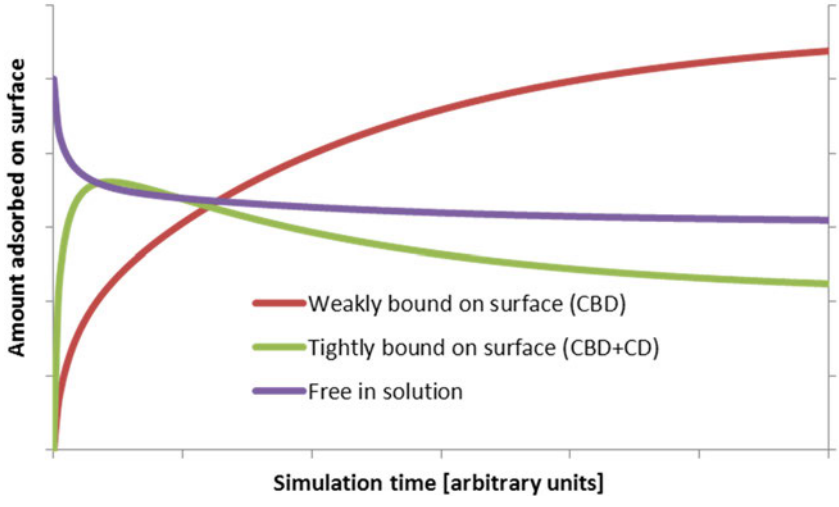

Figure 7.7. Hypothetical example of an almost constant level of total adsorbed enzyme despite changes in the population of adsorption sites. The example shows a greatly exaggerated effect.

7.6.4 Rate of adsorption

The rate of adsorption depends on the rate of mixing (Nidetzky *et al.,* 1994). Without complications, enzyme adsorption onto cellulose is believed to achieve equilibrium within approximately 60 minutes (Wang *et al.,* 2011). However, due to the

progressive enzymes sweeping across the surface of cellulose, the rate of adsorption will have to be fit for each substrate and enzyme cocktail combination.

7.6.5 Temperature dependency of dissociation constants

This chapter pertains to enzyme adsorption onto cellulose. The dissociation constants are affected by temperature (Heiss-Blanquet *et al.,* 2011). The enthalpies of dissociation according to Equation (7.6.1) was fitted to dissociation constants at 5°C (Kim & Hong, 2001), 30°C (Igarashi *et al.,* 2006; Igarashi *et al.,* 2011), and 45°C (Jäger *et al.,* 2010). The relatively large collection of Langmuir adsorption constants collected by Zhang and Lynd (2004) could not be used because they assumed only one type of binding sites.

$$\ln K_{T2} = \ln K_{T1} - \frac{\Delta H_{diss}}{R}\left(\frac{1}{T_2} - \frac{1}{T_1}\right) \tag{7.6.1}$$

The enthalpy of dissociation for the tighter binding site for Cel7A is probably somewhere between 43.6 kJ/mol and 63.6 k J/mol, more probably between 46 kJ/mol and 51 kJ/mol. The dissociation constant is likely to be between 0.09 µM and 0.31 µM. The dissociation enthalpy value ΔH_{diss} = 43.6 kJ/mol is consistent with a dissociation constant K_{diss} = 0.184 µM at 30°C. The above correlation is valid from 4°C to 45°C, although it can be used at even higher temperatures.

The enthalpy of dissociation for the weaker binding site for Cel7A was fitted to only two temperatures since no data was available for 45°C. The dissociation enthalpy value ΔH_{diss} = 75.8 kJ/mol is consistent with a dissociation constant K_{diss} = 2.3 µM at 30°C. The above correlation is valid from 4°C to 30°C, although it can probably be used at higher temperatures. With these parameters, the dissociation constant for the weaker binding site would be around 10 µM at 45°C. With typical enzyme dosages around 1 µM, this would not easily show up in adsorption measurements where the adsorption reaction is terminated with centrifugation.

According to Heiss-Blanquet *et al.* (2011), endoglucanase 1 (EG1) is desorbed in equal proportions to cellobiohydrolase 1 (CBH1) when going from 4°C to 45°C. Therefore, Cel7B was assumed to have the same biding enthalpies as Cel7A. Jäger *et al.* (2010) measured a dissociation constant K_{diss} = 2.2 µM at 45°C. The weaker binding site for Cel7B was assumed to have the same dissociation constant as Cel7A at 30°C.

Both enthalpies of dissociation for Cel6A were assumed identical to those of Cel7A. After Kim and Hong (2001), the dissociation constant for the weaker biding site is 13.6 nM at 5°C, and the dissociation constant for the tighter biding site is 119 nM at 5°C.

According to Heiss-Blanquet *et al.* (2011), 15% to 38 % less CBH1 was adsorbed at 45°C than at 4°C. The correlations made here estimate that the corresponding value is between 23% (tight binding site) and 85% (weak binding site). The decrease of adsorbed CBH2 at 45°C compared to 4°C was between 50% and

100% (Heiss-Blanquet *et al.,* 2011). Although the number of weak and strong binding sites are unknown in the Heiss-Blanquet case and CBH2 generally adsorbs weaker than CBH1, their results can be considered approximate validation of our dissociation parameters. It is also interesting to note that adding bovine serum albumin at 4°C has approximately the same effect as increasing the temperature (Heiss-Blanquet *et al.,* 2011), i.e. the weakly bound enzymes are desorbed.

Temperature should not affect the maximum number of adsorption sites as long as the enzyme remains soluble (Kim & Hong, 2001; Heiss-Blanquet *et al.,* 2011; Igarashi *et al.,* 2006; Igarashi *et al.,* 2011). The initial surface area for adsorption is much larger for cellulose III than for cellulose I (Igarashi *et al.,* 2011). The average surface area needs to be fitted for substrates coming from each type of pretreatment. Furthermore, the initial numbers of weak and tight binding sites need to be fitted.

7.6.6 pH dependency of dissociation constants

Binding of the wild-type Cel7A CBD is practically insensitive to pH (Linder *et al.,* 1999). The small dependence is roughly captured by a quadratic curve fit (7.6.2), where a = -0.008180854, b = 0.074401689, and c = 0.830836384. Cel7B and Cel6A were assumed to not have a pH-dependency.

$$rfExoProgAds = (a * pH + b) * pH + c \quad (7.6.2)$$

The bindings of Cel7B and Cel6A to cellulose were assumed to not have a pH-dependency.

7.6.7 Temperature dependency of adsorption rate

The temperature dependency of the adsorption rate measured at 4°C, 22°C and 30°C (Linder & Teeri, 1996) was fitted to an Arrhenius equation. The enthalpy of activation was 70 kJ/mol with k_{ref} = 0.028467938 at 30°C. All other adsorption rates were assumed to have the same activation enthalpy.

7.6.8 Inhibition by small molecules

Cellulose-binding domains have three tyrosine or tryptophan moieties spaced ideally for adsorbing onto three glucan repeating units. If the aromatic structure is capable of associating with glucan repeating units, then it is likely that it also associates with its smaller counterparts: glucose and cellobiose.

Small molecules such as glucose, cellobiose, and xylose may directly interact with the cellulose-binding domain of the enzyme, condense onto the surface of cellulose thereby reducing the number of effective adsorption sites, or they may slow down the movement of enzymes by increasing the apparent viscosity. This may mathematically be described as competitive inhibition parallel to enzyme adsorption onto cellulose.

7.7 Hydrolytic reactions

The saccharification of cellulose proceeds either through the cellobiose intermediate or directly to glucose (Figure 7.8).

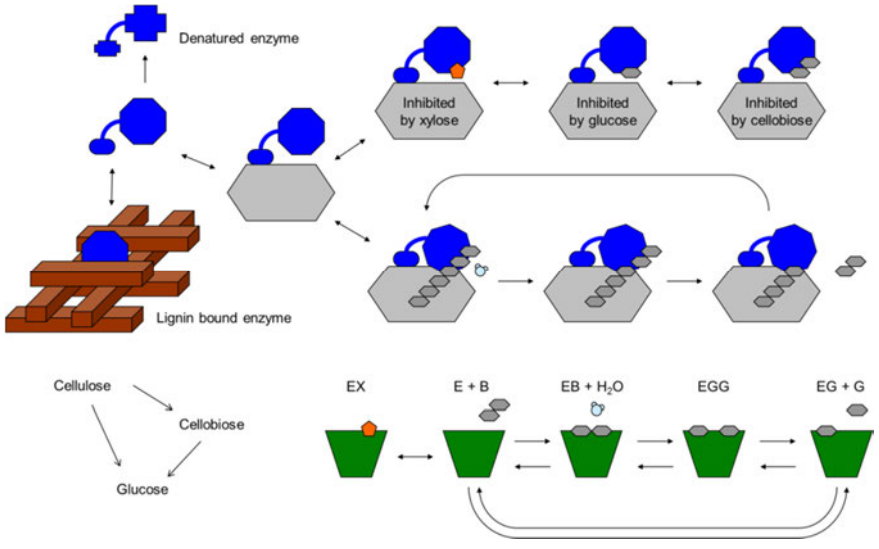

Figure 7.8. Primary enzymatic mechanisms in cellulose hydrolysis. The abbreviated notation refers to β-glucosidase enzyme (E), cellobiose (B), and glucose (G).

Once the enzyme has adsorbed onto cellulose, the active center needs to find a free chain-end. It is often presumed that small molecules competitively inhibit this step (Bansal *et al.,* 2009). After the hydrolytic step, the product molecule desorbs. While desorption is often assumed instantaneous (Bansal *et al.,* 2009), we assume it is in equilibrium. Since we have no direct evidence, i.e. measurements to fit the parameters, we assume that the binding constant is determined by the solubility of the product.

After product desorption, exoglucanases need to slide along the chain to the next position (Beckham *et al.,* 2011; Igarashi *et al.,* 2011). This progression step is unlikely to be a limiting factor, although traffic jamming on the surface of cellulose has been proposed in literature (Igarashi *et al.,* 2011). However, the rate is not necessarily strictly proportional to the free surface area. It actually tapers off and attains a constant value for large empty surfaces.

β-Glucosidases bind to cellobiose, whereupon a water molecule executes the hydrolytic reaction. The two product molecules are then desorbed one at a time. All reactions are reversible and limited by rate expressions.

Understanding the extent and mechanisms the appa rent inhibition of β-glucosidases could improve the basis for optimizing saccharification systems

(Bohlin *et al.,* 2010). The above scheme is superior to any overall Henri-Michaelis-Menten based kinetic expression regardless of the inhibition scheme used. The existence and necessity of both EG and EGG complexes make all simple forms of product inhibition schemes fail. Furthermore, the reversibility of the net reaction has been thoroughly documented (Andrić *et al.,* 2010a & 2010b).

7.7.1 Beta-glucosidase

The kinetic parameters for the inter-conversion of β-glucosidase and all its substrate-enzyme complexes were fitted. The parameters were selected such that the apparent equilibrium constant for the hydrolysis reaction connecting cellobiose and glucose was satisfied at 50°C. The apparent equilibrium constant at 50°C was calculated from thermodynamic parameters published by Tewari *et al.* (2008) and Goldberg and Tewari (1989). The reaction quotients at the end of many hydrolysis experiments are far smaller than the equilibrium constant. It also appears that the more enzyme is used, the higher the reaction quotient. We hypothesize that the existence of the EGG complex slow down the rate of hydrolysis to such an extent that equilibrium conversion is not reached within typical hydrolysis times.

The majority of the apparent kinetic constants were taken from literature studies focusing on this particular reaction (see below). In the case of β-glucosidase, Isothermal Titration Calorimetry has proven to be a convenient quantification tool for the rate of reaction (Tewari *et al.,* 2008; Bohlin *et al.,* 2010). Since the enthalpies of transglycosylation reactions are negligible, the enthalpy of hydrolysis is measured directly (Bohlin *et al.,* 2010). The method relies on measuring the enthalpy of reaction and thus avoids the problems with quantifying subtle changes in substrate and product concentrations. Jeoh *et al.* (2005) published kinetic parameters including substrate inhibition. Bohlin *et al.* (2010) studied the kinetics of the hydrolysis of cellobiose.

7.7.2 The progressive exoglucanase Cel7A

Cel7A activity is linearly proportional to the amount of enzyme adsorbed onto suitable reducing ends on the substrate (Nidetzky *et al.,* 1994). Furthermore, the movement requires some free surface area and therefore the rate is also proportional to the free surface area. However, since the reach of a single enzyme is limited, the amount of free surface quickly becomes irrelevant with increasing free surface area as long as the immediate vicinity of the enzyme is free. This is well captured by saturation kinetics.

The reducing end surface sites specific for progressive exoglucanases are assumed inaccessible to the other enzymes. The CBD of Cel7A may compete with Cel7B for other surface sites. Progressive exoglucanases tend to get stuck (see Chapter 1.9 for details). Once they pile up on one another, their combined pushing force may peel off chain segments (Igarashi *et al.,* 2011). This was modeled as

spontaneous release from the stuck state with a rate proportional to the surface concentration of stuck enzyme squared.

7.7.3 The non-progressive exoglucanase Cel6A

The non-progressive exoglucanase Cel6A adsorb onto the non-reducing end of cellobiose chains and sharpen up the fiber (Igarashi *et al.,* 2011) via a clipping action. The primary product is a cellobiose molecule although some glucose is also produced. See Chapter 1.7.8 for details on the product specificity.

The rate is proportional to the amount of Cel6A adsorbed onto suitable non-reducing ends. Cel6A may also interact near nicks, making them larger. When nicks are large enough, they count as free chain ends for progressive exoglucanases.

The CBD of Cel6A may also compete with Cel7B for surface sites. Although there are some reports that other enzymes compete with Cel6A for the non-reducing ends (Nidetzky *et al.,* 1994), this effect was omitted.

7.7.4 The endoglucanase Cel7B

The endoglucanase Cel7B adsorb onto the surface of cellulose and use a water molecule to hydrolyze a bond along the cellulose chain. The primary product is a nick in the chain. However, if two nicks are close enough to each other, a glucose or cellobiose molecule may detach from the surface. The rate is proportional to the amount of Cel7B tightly adsorbed onto the surface.

The action of the endoglucanase enzyme may set free stuck enzymes by making the obstacles smaller.

Endoglucanase slow down due to surface crowding (Nidetzky *et al.,* 1994), but this is a consequence of competition for adsorption spots. Only when the catalytic domain is adsorbed onto the surface, the enzyme is active. When the enzyme concentration increases, the preferred lattice structure has less catalytic domains and more cellulose binding domains. See Chapter 1.6.3 for details on the binding sites.

7.7.5 Temperature dependency

In the models by Kadam *et al.* (2004) and Zheng *et al.* (2009) the apparent temperature dependency of reaction kinetics was approximated without denaturation. By explicitly including temperature-dependent denaturation and after refitting the enthalpy of activation, the present model is valid for a wider range of temperatures.

Measurements at optimal pH and at different temperatures are often reported in publications of enzyme characterization. The data series will show an exponential temperature dependency, which usually taper off when measurement uncertainties become dominant in the high-rate limit. The temperature dependency is hardly ever anything other than an Arrhenius parameter.

Arrhenius-like activation energies were fitted to the mostly exponential parts of the temperature-dependent rates. The value 38.0 kJ/mol for the progressive exoglucanase Cel7A was obtained by fit to Figure 3C in Karboune *et al.* (2008) (*P. funiculosum* cellobiohydrolase). The value 48.4 kJ/mol for the non-progressive exoglucanase Cel6A was obtained by fit to data from Figure 1 in Baker *et al.* (1992) (*T. reesei* cellobiohydrolase II, avicel substrate). The value 38.0 kJ/mol for the endoglucanase Cel7B was obtained by fit to Figure 3A in Karboune *et al.* (2008) (*P. funiculosum* endo-1,4-β-D-glucanase). The value 52.5 kJ/mol for β-glucosidase was obtained by fit to Figure 3B in Karboune *et al.* (2008) (*P. funiculosum* β-glucosidase). It is a mere coincidence that two of the values are almost identical.

7.7.6 pH dependency

Hydrolytic enzymes usually fall into the enolase family, where the active center typically contains a catalytic triad. Regardless if the mechanism is inverting or non-inverting, it occurs by several shifts in electron density between this triad of amino acid side chains. The rate of such electron shifts is inversely proportional to the difference in the acid dissociations constants of the two participating residues. Because of this mechanism, the pH-optimum is quite limited.

The pH-dependency of hydrolytic enzymes is often visually described as a bell-shaped curve. Classical bell-shaped curves (7.7.1) describe two acid dissociation constants and adequately capture the enzyme activity near the pH-optimum.

$$adsC_{\text{active}} = \frac{adsC_{\text{tot}}}{\frac{10^{-pH}}{10^{-pK_{a1}}} + 1 + \frac{10^{-pK_{a2}}}{10^{-pH}}} \tag{7.7.1}$$

An enzyme having two acids with different pKa values can be in three states: both acids protonated, one acid protonated, and both acids deprotonated. The concentration of the intermediate constituent has the characteristic shape of a bell. If only this constituent is capable of catalyzing a reaction, the reaction rate and therefore the extent of reaction in a fixed-time assay will have the same shape.

These acids can be assumed to be essential to at least one limiting step. If all three states are capable of catalyzing the reaction and they have different turnover numbers, then the overall rate is described by the sum. Thus, the enzymatic activity may be different in each state. Depending on the relative magnitude of the activities, the bell-shaped curve could be standing on a staircase, i.e. the enzyme activity not going to zero at both ends. A third acid is typically seen as a shoulder if the limiting step changes between states.

The pH-activity curve is not only pKa-sensitive but also slope-sensitive and the slopes are different for the acidic and basic branches (Maurelli *et al.,* 2008). The curvature is rarely a perfect sigmoidal shape, but stretched or compressed in some way. Simple stretching or compression comes from having a population of enzymes, each with slightly different chemical environment.

The distribution of electron density will be different in and around the active centers of the three constituents. Reaction mechanisms are essentially only electron density moving from one place to another. Therefore, one may expect that the rate-limiting steps may be different for the three environments. By itself, this may not be important, but it makes a difference for the acid-base equilibria.

The local electron density around ionizable groups affects their apparent pKa, which complicates the modeling. The rates of acid-base equilibria compete directly with the rate-limiting step of the enzyme-catalyzed reaction. This leads to the concept of characteristic times for each rate-limiting step. Furthermore, each rate-limiting mechanism may be affected differently by foreign ions and Donnan partitioning. Consequently, the effective concentration experienced by a specific rate-limited mechanism is different from the concentration it would be at perfect acid-base equilibrium. In practice, it means that the concentrations of constituents in Equation (7.7.6.1) can be construed as separate acid-base equilibria. It also implies that symmetry is not a required property. Some experimental pH-dependencies are symmetric (Ruttersmith & Daniel, 1991) and some are not (Maurelli *et al.,* 2008). Many are hard to distinguish (An *et al.,* 2004; Becker *et al.,* 2001) and some depend on the presence of ions (Huang *et al.,* 1988).

If denaturation can be ignored, the slope is inversely proportional to the time constant of the event step that particular amino acid is involved in. A steep curve and thus a large time constant indicate an event that lasts longer than the rate-limiting step. The slope-determining factors describe e.g. the effects of foreign ions and Donnan partitioning.

While a detailed thermodynamic-kinetic model could have been constructed for describing these distorted bell-shaped curves, we went with a computationally easier approach. The engineering approach was to reproduce the measured curve as closely as possible but without dwelling too much on the detailed mechanism behind it. The parameters used for replicating the curves are not intended to hold any physical significance since they contain multiple effects.

The parameters for the progressive exoglucanase Cel7A was obtained by fitting the chosen bell-shaped curve (7.7.2) to data from Figure 1C in Karboune *et al.* (2008) (*P. funiculosum* cellobiohydrolase). The parameters for the non-progressive exoglucanase Cel6A was obtained by fit to data from Figure 2 in Koivula *et al.* (2002) (*T. reesei* Cel6A). The parameters for the endoglucanase Cel7B was obtained by fit to Figure 4A in Kwon *et al.* (1999) (*T. viride* endoglucanase I). The parameters for β-glucosidase were obtained by fit to Figure 1B in Karboune *et al.* (2008) (*P. funiculosum* β-glucosidase). The curves corresponding to these parameters are plotted in Figure 7.9.

$$act_{\mathrm{rel}} = \left(\frac{1}{scalefactor}\right)\left(1-\left(1+10^{\frac{pH-pHa1}{taufactor1}}\right)^{-1}-\left(1+10^{-\frac{pH-pHa2}{taufactor2}}\right)^{-1}\right) \qquad (7.7.2)$$

The *scalefactor* is intended solely for normalization. The *taufactors* affect the slopes and the pK_a values affect the width of the bell.

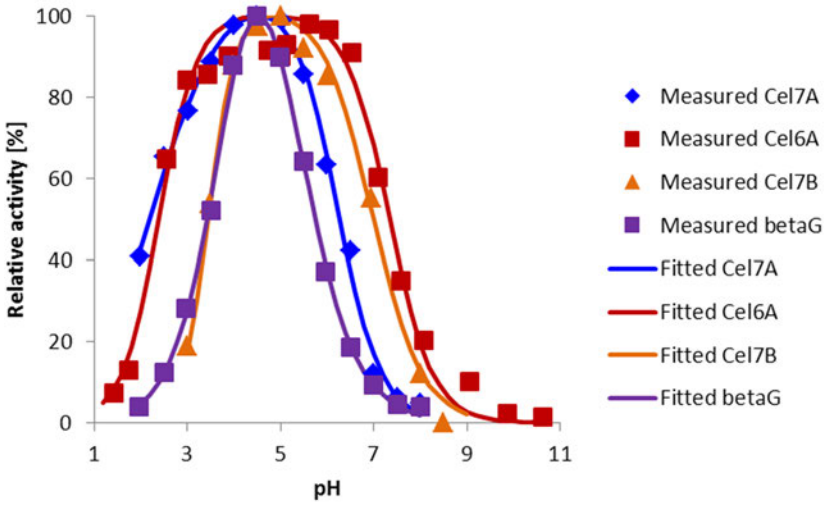

Figure 7.9. pH dependency of hydrolytic rates for Cel7A, Cel6A, Cel7B, and β-glucosidase.

7.7.7 Effect of ions

Ions affect the rate of hydrolysis in the entire pH and temperature range of hydrolytic enzymes (Huang *et al.,* 1988). Both anions and cations display a double binding relationship between activity and the concentration of the ion (An *et al.,* 2004; Huang *et al.,* 1988). Fortunately, the amount of ions coming along with the biomass is in most cases so low that only the first binding interaction needs to be modeled. If there are additional ions from pretreatment or added with the pH buffer, these concentrations may exceed the threshold and the second binding event is needed.

Since the effect of different cations is not additive (Huang *et al.,* 1988), their binding to the active center must be competitive. The combination of size-exclusion and charge-exclusion could explain the competition and the total loss of activity induced by heavy metals. Heavy metals also catalyze the denaturation of enzymes.

The two most common pH-buffers are acetate and citrate buffers. A buffer is theoretically not needed since the hydrolysis reaction does not create any new acids. However, the problem with an unbuffered solution is that the precious enzymes may denature prematurely or perform poorly due to less than optimal pH values. The buffer anions also bind metal ions. Depending on the concentration of metal ions, the effect can be either positive or negative.

Some versions of our hydrolysis models implement a thermodynamic solution model describing the aqueous solution, including buffer action, ash release, and some complexation. These implementations are based on thermochemical calculation software (Koukkari, 2009; Blomberg & Koukkari, 2011; Koukkari *et al.,* 2011; Blomberg & Koukkari, 2009).

7.7.8 Product specificities

The product specificities of hydrolytic enzymes probably depend on a lot of factors and are likely to change during the course of hydrolysis. However, a fixed product profile was assumed for each of the enzymes.

The average product specificity of the progressive exoglucanase (Cel7A) and endoglucanase (Cel7B) were adopted from Eriksson *et al.* (2002). The progressive exoglucanase Cel7A make about 80 mol-% cellobiose and 20 mol-% glucose. The endoglucanase side reaction produced on average 55 mol-% cellobiose and 45 mol-% glucose.

The pH-dependent product profile of non-progressive exoglucanase (Cel6A) was fitted to the two data points at pH 5 and pH 7 in (Wohlfahrt *et al.,* 2003). The enzyme Cel6A strongly favors cellobiose over glucose.

The enzyme β-glucosidase produces glucose from cellobiose and thus has a trivial product profile.

7.7.9 Inhibition by small molecules

The active center of all hydrolytic cellulases has a large number of hydrogen bonds interacting with the substrate. Thus, it can be expected that almost all sugars present some inhibitory effect. Any mutation will probably reduce the inhibitory effect of one substance but increase it for another. Compounds very similar to sugars, such as 5-HMF and furfural, may also inhibit the enzyme by competitive inhibition. Furthermore, furans tend to chelate metal cations, which may or may not be required for enzymatic activity.

Fermentation products like ethanol and glycerol usually act as co-solvents, thereby reducing the force keeping the hydrophobic regions out of water. While sometimes classified as competitive inhibition, the effect of ethanol might be better described by reversible denaturation. See Chapter 7.11 for more details.

7.8 Product desorption

Product desorption is a key stage in the hydrolysis of cellulose and cellobiose. It is rarely instantaneous except at very low substrate loadings. It requires not only solvent (water), but also adequate mixing/diffusion. The desorption constants were approximated by the solubility constants due to the lack of measurements specifically quantifying product desorption.

In the case of β-glucosidase, the two-step desorption of products imply vastly nonlinear inhibition that can not be adequately captured by a single competitive, non-competitive, or mixed inhibition term. Modeling the nonlinear nature requires two terms instead of one.

The temperature dependent solubility of glucose, xylose, and cellobiose was calculated from the thermodynamic data published by (Tewari *et al.,* 2008; Gold-

berg and Tewari, 1989). At temperatures below 53°C, glucose monohydrate is the more stable solid, thus severely limiting the solubility of glucose. Whether or not this remains an accurate analogy for product desorption depend on the existence of a single water molecule. The higher the temperature, the more glucose may dissolve, and product desorption is a smaller problem. As it turns out, only the solubility of cellobiose is severely limiting the overall conversion of cellulose into glucose.

The solubilities reported by Gray *et al.* (2003) seemed slightly better at 20°C. Therefore, the solubilities predicted by the thermodynamic parameters were adjusted by a factor such that they coincided with the solubility measurements performed by Gray *et al. (*2003) at 20°C. This scaling was needed only for glucose and xylose.

7.8.1 Temperature dependency of dissociation constants

The temperature dependency of the dissociation constants, *i.e.* the solubilities, were calculated from the standard enthalpy of formation, the standard entropy of formation, and the heat capacity of each of the constituents involved.

The same data allows the calculation of the equilibrium constant between cellobiose and glucose. The equilibrium constant reduces with increased temperature. Glucose is thus favored at low temperatures while cellobiose is more favored at high temperatures. At 50°C and high solid loading, the equilibrium extent may indeed be limiting the amount of glucose compared to the amount of cellobiose. If the same scaling for glucose is allowed as in the above solubility case, then the equilibrium constant is a real concern. The difficult is in determining the (thermodynamic) activities of glucose and cellobiose in this highly non-ideal mixture of solids, ash, ions, enzymes, and soluble organic compounds.

7.9 Traffic jamming

The linear progression of Cel7A along the length of the cellulose fiber can be described with an analogy of traffic lanes. Each enzyme requires a certain amount of room, more room than a single lane would give them. While this is like modeling an extra wide transport, every vehicle is an extra wide transport occupying neighboring lanes. Traffic jams are therefore much more frequent. If the amount of Cel7A on the surface is too high, they will interact with each other and slow down the overall rate of hydrolysis (Igarashi *et al.,* 2011; Palonen *et al.,* 2004; Suchy *et al.,* 2011). Specifically, when buffer rinse is introduced in a system with high enzyme loading, some of the enzyme overcrowding is removed and rapid hydrolysis follow.

The enzymes can also get stuck if lignin or another cellulose fiber lay across the lane a progressive enzyme attempt to traverse.

7.10 Cellulose surface area and thickness

The hydrolytic enzymes Cel7A, Cel6A, and Cel7B display synergy both when acting simultaneously and sequentially (Zhang & Lynd, 2004; Bansal *et al.,* 2009). An elaborate construction of different adsorption surface sites and mechanistic actions was designed to reproduce some of the known relationships. Since the degrees of synergy depend not only on the enzymes but on the substrate, it has been left to the user to fit appropriate parameters for their particular system.

Heiss-Blanquet *et al.* (2011) BSA binding studies and Igarashi *et al.* (2011) atomic force microscopy (AFM) imaging indicate that Cel6A chip away and sharpen the non-reducing ends of cellulose crystals, thereby reducing the total number of adsorption events linearly with time. The same studies show that the adsorption of endoglucanases does not change at first and then reduce to a small value. Similarly, it seems that Cel7A adsorption increase at first and then decrease to effectively zero.

Exoglucanases increase the surface concentration of prime endoglucanase spots by removing longer chain fragments. Similarly, endoglucanases increase the surface concentration of prime exoglucanase spots by introducing nicks. Endoglucanases and the non-progressive exoglucanase may then enlarge nicks into chain-ends. According to Table 2 in (Palonen *et al.*, 2004), the surface will first run out of exoglucanase spots, presumably because they require long cellulose chains.

As long as the computationally convenient cellulose fiber thickness is larger than a single layer of cellulose strands, the consumption of a chain-end in the reducing end produces another. The progressive exoglucanase will consume the chain-end if it its origin is a nick. Similarly, the non-progressive exoglucanase action reduces the number of Cel6A-specific surface sites only when the thickness is low. Similar constructions have been proposed since Converse and Optekar (1993).

In the end we ended up with 5 surface sites. The three major enzymes modeled adsorb to these sites according to Table 7.1. Once adsorbed, there are six enzyme activities to model. The reaction stoichiometries depend a little on the amount of cellulose according to Table 7.2 and Table 7.3. The possible cell entries are “substrate”, “product”, “increases”, and “decreases”. The first two indicate that their stoichiometric coefficient is equivalent to unity. The specific term “substrate” also indicates that it is absolutely necessary for the reaction to proceed. If a quantity “decreases”, it means that it is only partially consumed e.g. because the previous exposes another from beneath, but one that is not as accessible. The “increasing” keyword is used when there is a finite nonzero probability of something other than the main function occurring immediately after a reaction step.

Table 7.1. The three enzymes (the progressive exoglucanase Cel7A, the non-progressive exoglucanase Cel6A, and the endoglucanase Cel7B) adsorb selectively to the five types of surface sites. The cell entry state "active" if the enzyme adsorb with both CBD and CD.

	Cel7A	Cel6A	Cel7B	Site description
A_prog	active			Reducing end
A_chip		active		Non-reducing end
A_chain	active			Loose chain on surface
A_nick		active	active	Crystal defect
A_endo	inactive	inactive	active	Crystal surface
A_nick		active	active	Crystal defect
A_endo	inactive	inactive	active	Crystal surface

Table 7.2. The consequences of each of the six adsorbed enzyme activities when the crystal is thicker than a single layer of cellulose strands. See text for explanations on cell entries.

	Cel7A on A_prog	Cel7A on A_chain	Cel6A on A_chip	Cel6A on A_nick	Cel7B on A_nick	Cel7B on A_endo
A_prog						
A_chip						
V_chip			decreases			
A_chain		decreases		increases	increases	
A_nick				decreases	decreases	product
A_endo	substrate and product	substrate and product		decreases	decreases	substrate
Products	product	product	product	product	product	
Thickness	decreases	decreases			decreases	decreases
Stuck Cel7A	increases	increases			decreases	decreases

Table 7.3. The consequences of each of the six adsorbed enzyme activities when the crystal is thinner than a single layer of cellulose strands. See text for explanations on cell entries.

	Cel7A on A_prog	**Cel7A on A_chain**	**Cel6A on A_chip**	**Cel6A on A_nick**	**Cel7B on A_nick**	**Cel7B on A_endo**
A_prog	decreases					
A_chip			substrate			
V_chip						
A_chain		decreases				
A_nick				substrate	substrate	
A_endo	substrate	substrate	substrate	substrate	substrate	substrate
Products	product	product	product	product	product	product

The progressive exoglucanase Cel7A peels the surface like a cheese-scraper. Each layer peeled away reveals another just like it from beneath until the thickness runs out. Therefore, the surface area (Aprog) will not change due to the action of Cel7A in Table 7.2.

The non-progressive exoglucanase Cel6A sharpens the non-reducing end one chip at a time (Igarashi *et al.,* 2011). It will consume a triangle in the perpendicular cross-section of the cellulose bundle. The volume of the triangle is determined by the original thickness and the original cellulose surface area. Therefore, the V_chip volume will only reduce during hydrolysis until there is no more chipping that can be done. The adsorption area A_chip will neither increase nor decrease due to the chipping action. It will at most be slightly moved sideways. The characteristics of each enzyme modeled are shown in Figure 7.10.

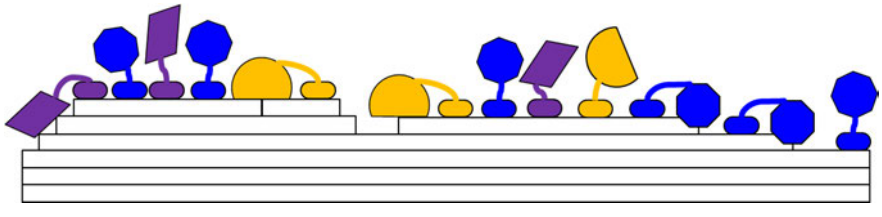

Figure 7.10. Perpendicular cross-section of a cellulose bundle. The white bars represent continuous cellulose chains in layers, the blue octagons are Cel7A, the yellow spheres are Cel7B, and the purple parallelograms are Cel6A. The leftmost endoglucanase has just made a nick in the chain underneath.

7.11 Ethanol inhibition

Ethanol is known to affect the rate of catalysis (Bezerra & Dias, 2005). However, the exact mechanism is sometimes uncertain. Ethanol acts like a co-solvent dis-

solving both small molecules and enzymes. Which of the effects is stronger at any particular instance is unknown. An ad hoc modulation based on data from Chen and Jin (2006) was included.

7.12 Conclusions

The mathematical model presented here describes the enzymatic hydrolysis of dilute acid pretreated lignocellulosic biomass. The model captures the trends and approximate magnitudes of effects relevant to the hydrolysis of cellulose.

This engineering model has the broadest operation range publicly known to date. The model includes effects of pH, temperature, solid loading, enzyme dosage, and process time. Since it is not a correlation model, the outcome can be simulated with or without select phenomena. Therefore, the relative importance of different phenomena can be investigated. The select phenomena include activity of water, instantaneous product desorption, surface crowding, enzyme adsorption onto lignin, temperature effects, pH effects, enzyme denaturation, and water retention value. Furthermore, parameters such as temperature, enzyme amounts, pH, and solid loading can be varied to a much greater extent than for correlation models.

The mechanistic action of the enzymes is described in high detail for an engineering model. The inclusion of phenomena like activity of water, desorption of product, and surface-movement of enzymes is exceptional among engineering models. Furthermore, the model contains rare parameters like water retention value, free surface area, and reversible kinetics for β-glucosidase. This may be the first model to describe such a variety of stalled enzyme complexes.

There are some important limitations to this model. The structure of the biomass is not included. Therefore, this model is quite specific to dilute acid pretreated straw-like lignocellulosic biomass. The enhancing effect at low concentrations and the detrimental effect at high concentrations of soluble ions are not included. Therefore, the precise influence of ash composition and the choice of buffer are not captured by the model. Enzyme denaturation due to unusual reduction potentials is not included. The kinetics is quite specific to Celluclast 1.5 and Novozyme 188.

The ionic strength and the interactions of specific ions with the enzymes have been omitted from this model. Enzyme activities are known to be ion-specific. It is also conceivable that a proportion of the ash-content of the raw material is retained after pretreatment. These inorganic ions and molecules may then affect the rate of hydrolysis in a matter not captured by the current model.

8. Fermentation

Bioreactor is a key unit operation in biorefinery processes. The hydrolyzed sugars are often converted to ethanol or other products by fermentation. Most of today's models for bioreactors are empirical data-based models and thus their operation window is rather limited and case specific. In this project, we used metabolic networks of the micro-organism as a basis of the bioreactor model. Metabolic networks present the reaction stoichiometries within the network and thus form a mechanism for the model. These so called dynamic flux balance models (DFBA) have been seen as a great opportunity for designing and optimizing fermentation processes. We have developed an efficient parameter estimation method for DFBA models, which is a key step in using the models in process engineering applications. We also used these models to estimate the isobutanol production by metabolically modified yeast.

8.1 Extracellular kinetics

The starting point for most of our fermentation models is a set of measured time profiles for extracellular metabolites and the evolution of carbon dioxide. Typically, the concentrations of acetic acid, ethanol, glucose, glycerol, xylose, fructose, and xylitol are analyzed by high-pressure liquid chromatography (HPLC). Minor acids like succinic acid, lactic acid, and malic acid may also be measured. The amount of biomass is often followed with both optical density measurements (OD) and as cell dry weight (CDW). The temperature and weight of the fermentation vessel are sometimes recorded continuously. The amount of acid or base required to keep the pH constant is routinely recorded. Dissolved oxygen and the oxygen difference in inlet and outlet gas streams can also be measured.

The precise protocol for fitting a set of ordinary differential equations (ODE) to describe the measurements is outside the scope of this text. However, rest assured that such fits can be made, especially with a little insight to the structure of the dependencies. The measurements rarely fit well to perfectly mass-balanced differential equations. In our experience, it is best to deal with such issues in later analyses rather than here in the first step.

An example of yeast growing on glucose is presented in Figure 8.1. Several distinct physiological states can be identified. First, the yeast grows on glucose and

produces ethanol, acetate, and glycerol. When the glucose runs out, there is an adaptation period. After the adaptation period, the yeast consumes acetate, ethanol, and glycerol. The acetate is depleted first. Shortly after, the glycerol is depleted. Finally, also the ethanol is depleted.

Yeast *Saccharomyces cerevisiae* strain CEN.PK113-7D was cultivated under aerobic conditions on 20 g/L D-glucose on Delft minimal mineral medium in a 5-liter B.Braun Biotech bioreactor at 30°C at pH 5 controlled using 2 M NaOH (Tohmola *et al.*, 2011). Extracellular D-glucose, acetate, glycerol and ethanol were measured automatically at 10-minute intervals using a Medicel on-line HPLC with cross-flow filtration for obtaining a cell-free sample. An HPX-87H fast acid column was used for analyte separation and refractive index detection was utilized for quantification. Cell density was measured from the same samples taken for the on-line HPLC using optical Trucell probe. Carbon dioxide and oxygen in the off-gas were measured using Bluesens probes.

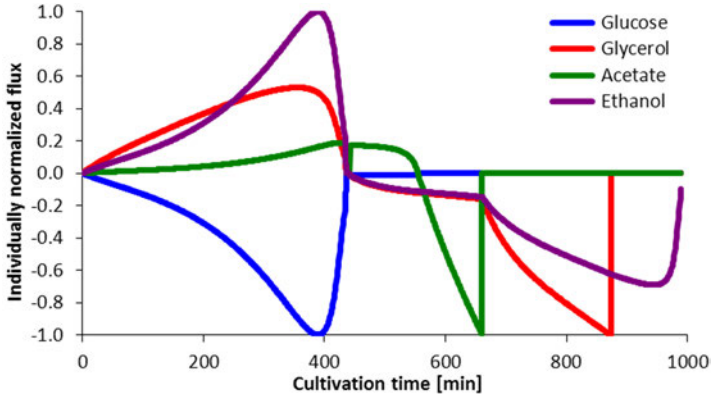

Figure 8.1. An example of yeast growth kinetics fitted to data published by Tohmola *et al*. (2011). The individually scaled boundary fluxes versus time accentuate the observed physiological states. 'Growth on glucose' ends after around 420 minutes. Although ethanol and glycerol are consumed during the 'adaptation period' from 420 to 550 and during 'growth on acetate' from 550 to 660 minutes, the period from 660 to 870 minutes is considered 'growth on glycerol' and the period from 870 to 1000 minutes is considered 'growth on ethanol'.

If the evaporation rate of ethanol is measured separately for the specific reactor used and the operating conditions employed, a correction for the loss of ethanol via the purge gas can be made.

8.2 Flux balance analyses

The primary purpose of these flux balance analyses was to extend the size of the system of differential equations. The known fluxes were calculated from the fitted

differential equations once the state vector was given. The remaining boundary fluxes were calculated based on flux balance, i.e. material conservation, and the mathematically convenient assumption of maximal biomass yield. Flux balance calculations were done separately for each time point in the series. The networks iND750 (Duarte *et al.,* 2004), iMM904 (Herrgard *et al.,* 2008), and a modification of the network in Pitkänen (2005) were used.

Depending on the scope of the network, the fluxes for H^+, $NH^{4+,}$ $HPO4^{2-}$, $SO4^{2-}$, etc. crossing the boundary of the cell are being calculated. These boundary fluxes are sensitive to the stoichiometry of the biomass-forming reaction. Both the known and the calculated boundary fluxes affect the extracellular solution. The pH of the extracellular solution would quickly drop if there were no pH controller. By knowing the hydrogen atom balanced fluxes across the boundary of the cell, the base consumption to keep the extracellular pH constant can be estimated. Figure 8.2 compares the base consumption estimates for iND750 and iMM904 to the experimental values. The figure also compares the biomass predictions for each of the considered metabolic networks.

For the inoculums, cells grown on an YPD-plate were taken into 25 ml Delft medium in 100 ml Erlenmeyer-flasks and incubated in a plate shaker (30°C, 150 rpm) over night. The cell broth was moved into 250 ml Erlenmeyer-flasks with 75 ml Delft medium and incubated in a plate shaker (30°C, 150 rpm) for 4–6 hours. At this point the cell density (OD 600) in the inoculums was about 3. An adequate amount of the cell broth was centrifuged (4°C, 2000 rpm, 5 min) and the cells were re-suspended into 40 ml of the growth media lacking the sources of carbon.

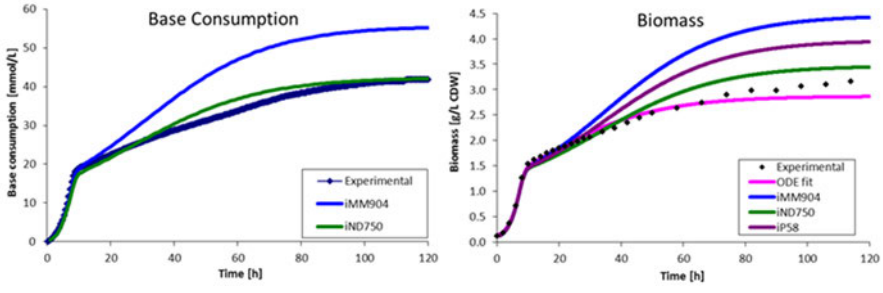

Figure 8.2. Comparison of experimental data and simulation based on genome-wide metabolic model for base consumption and biomass accumulation. *Saccharomyces cerevisiae* was grown anaerobically on glucose and xylose.

8.3 Minimal internal fluxes

Intracellular flux patterns are alternate, distinct, yet equally optimal flux-balance solutions (Mahadevan & Schilling, 2003) obtained using only external metabolite uptake and secretion measurements. Their analysis may provide useful biological insight (Baughman *et al.,* 2011). The robustness of a particular flux distribution at a single time point can be evaluated using flux variability analysis or modifications thereof.

The success, robustness, and reliability of dynamic flux balance calculations can be evaluated by exploring the feasible flux space at the optima. Even with two metabolic objectives, the feasible space, i.e. the number of alternate equally optimal flux distributions, is rarely a singular point when using genome scale metabolic networks. Assuming either transcriptional or translational constriction, the feasible space can be further reduced. The so called minimal internal fluxes are alternate solutions to a flux balance problem where the fluxes are additionally limited by either transcriptional or translational constriction.

Minimal flux patterns may be determined either by minimizing the number of non-zero reactions or by minimizing the sum of flux magnitudes. Assuming transcriptional constriction, the number of different enzymes is as small as possible since once and RNA strand has been created, any number of enzymes may be produced. Assuming translational constriction implies that the total amount of enzyme is limiting. Reality may not fit either extreme, but these are the simplifications needed for making the mathematical description computationally tractable.

The algorithms by Lee *et al.* (2000) and Murabito *et al.* (2009) are quite effective in producing all alternative patterns for a specific optimal metabolic objective. Minor variations of these algorithms find the minimal number of internal fluxes. Since it has been proposed that the fluxes of most enzymatically catalyzed reactions are limited by translation rather than transcription, it makes sense that the cell would minimize the total amount of protein required. While this amount can be estimated based on catalytic efficiencies (Shlomi *et al.,* 2011), minimizing the sum of absolute flux magnitudes is proposed as a substitute. Unlike the total flux objective by Holzhütter (2004), this objective avoids using equilibrium constants, which are a significant source of uncertainty due to poorly predictable Gibbs energies in genome wide networks.

For each time point, the limits of select boundary fluxes were set by the measured changes in the extracellular composition. Biomass yield was optimized by flux balance analysis on a genome scale metabolic network. In a sequential optimization step, the sum of fluxes was minimized while combinatorially varying the allowable flux patterns for a fixed biomass flux value. The proposition of flux patterns was recursive like in a branch and cut algorithm. Surprisingly, only a single level of recursion was required, making the algorithm fairly time-efficient. The second optimization uses a network with only non-negative fluxes. Thus, the algorithm finds all alternative flux distributions for the minimal sum of absolute flux magnitudes.

The differential equations relating to Figure 8.1 were used as time-dependent boundary constraints in dynamic flux balance calculations with the iMM904 network. Alternative flux distribution sets were obtained for each time point. Each set of minimal internal fluxes contain 5 to 10 flux distributions. Five physiological states were identified by comparing minimal flux patterns for subsequent time points. Figure 8.3 shows which time points share patterns. Both rows and columns represent time points equally spaced 30 min apart. Growth states are highlighted with a blue background color. The most notable difference to Figure 8.1 is that the ‘adaptation period’ begins roughly half an hour before glucose is completely depleted.

A highly ordered sequence of changes in most pathway feeds occurs during 'acetate adaptation'. The alternate carbon metabolism runs hyperactive. For each pathway lead, multiple sequential reactions exchange in 'Glycine and Serine Metabolism', 'Oxidative Phosphorylation', 'Phospholipid Biosynthesis', 'Purine and Pyrimidine Biosynthesis', 'Pyruvate Metabolism', 'Sterol Metabolism', 'Fatty Acid Biosynthesis', 'Tyrosine, Tryptophan, and Phenylalanine Metabolism', 'Valine, Leucine, and Isoleucine Metabolism', and in the 'Nucleotide Salvage Pathway'. Minor rerouting occurs for Alanine, Aspartate, Threonine, Lysine, and Glutamate syntheses.

Just prior to 'growth on acetate', glycolysis is reversed. When acetate runs out just prior to 'growth on glycerol', NAD/NADP cofactor specificities change, ATPase is activated, and TCA cycling is turned on. When glycerol uptake ceases, the source of glycerol for lipid synthesis is switched from uptake to synthesis.

	Cultivation time [min]																															
Cultivation time [min]	10	40	70	100	130	160	190	220	250	280	310	340	370	400	430	460	490	520	550	580	610	640	670	700	730	760	790	820	850	880	910	940	970
10	6	5	5	4	4	4	4	4	4	4	4			3																			
40	5	6	6	3	3	3	3	3	3	3	3			2																			
70	5	6	6	3	3	3	3	3	3	3	3			2																			
100	4	3	3	6	6	6	6	6	6	6	6			4																			
130	4	3	3	6	6	6	6	6	6	6	6			4																			
160	4	3	3	6	6	6	6	6	6	6	6			4																			
190	4	3	3	6	6	6	6	6	6	6	6			4																			
220	4	3	3	6	6	6	6	6	6	6	6			4																			
250	4	3	3	6	6	6	6	6	6	6	6			4																			
280	4	3	3	6	6	6	6	6	6	6	6			4																			
310	4	3	3	6	6	6	6	6	6	6	6			4																			
340												6																					
370													6																				
400	3	2	2	4	4	4	4	4	4	4	4			5																			
430															9																		
460																7																	
490																	8																
520																		8															
550																			8	1	1	2											
580																			1	8	4	3											
610																			1	4	7	1											
640																			2	3	1	7											
670																							10	5	2	2	3	2	3				
700																							5	10	5	5	6	5	6				
730																							2	5	8	7	7	7	6				
760																							2	5	7	8	8	7	6				
790																							3	6	7	8	9	7	6				
820																							2	5	7	7	7	8	7				
850																							3	6	6	6	6	7	8				
880																														10	10	10	10
910																														10	10	10	10
940																														10	10	10	10
970																														10	10	10	10

Figure 8.3. Flux pattern versus flux pattern matches for time points during the cultivation. The blue boxes show distinct physiological states. The experimental data of the aerobic yeast *Saccharomyces cerevisiae* cultivation demonstrates a Crabtree effect of ethanol, glycerol and acetate being produced from glucose and then consumed after glucose depletion and diauxic shift.

The measured (Tohmola *et al.*, 2011) and calculated biomass formation rates differ at the end of growth on glucose and during growth on acetate (Figure 8.4).

The former may be attributable to volume restrictions, i.e. the cell having an upper limit for the total amount of protein present (Baughman *et al.*, 2011). Except for the anomalous peak top, there are two continuously accessible patterns during growth on glucose. One pattern has 299 nonzero fluxes, the other 300. The latter discrepancy in Figure 8.4 is probably due to erroneous ATP cost of acetate uptake or the use of an alternate acetate uptake path not included in the network.

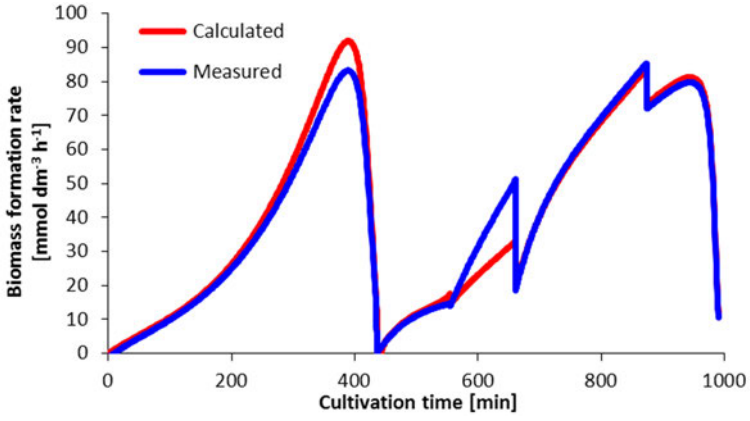

Figure 8.4. Measured and calculated biomass formation rates differ at the end of 'growth on glucose' and during 'growth on acetate'.

Figure 8.5 and Figure 8.6 show how the two objectives for calculating the minimal internal fluxes compare in each other's eyes. Figure 8.5 show the expected consequence of trying to duke it out with as few enzymes as possible but later having to create new ones anyway. Figure 8.6 on the other hand is quite interesting.

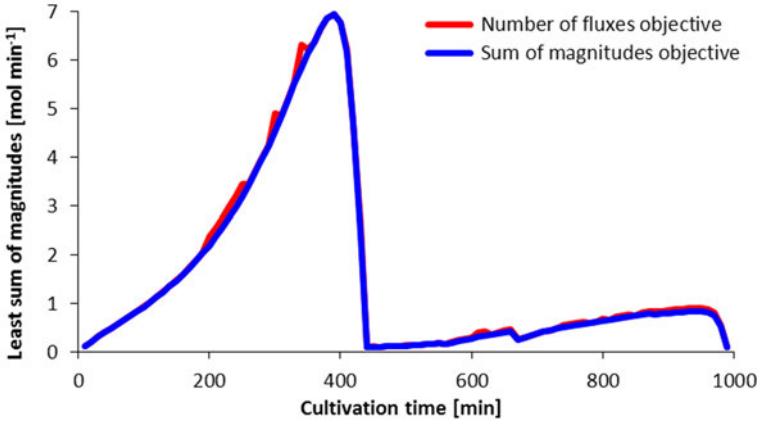

Figure 8.5. This comparison of the two objectives for calculating minimal internal fluxes shows the sum of magnitudes with respect to cultivation time.

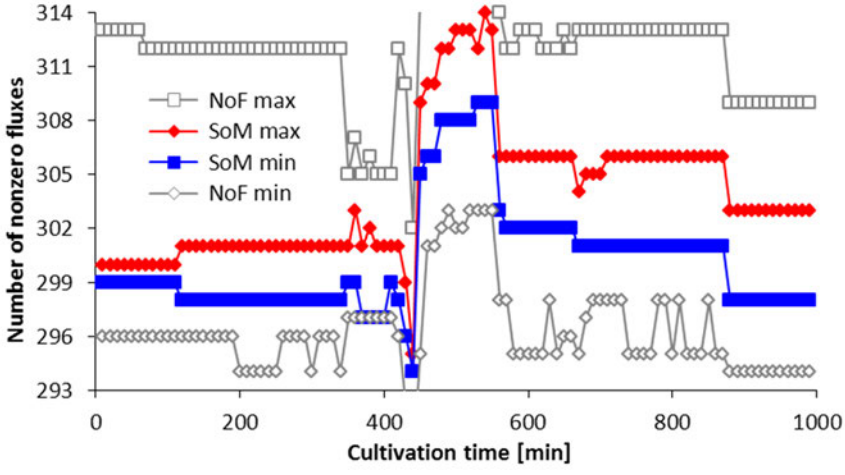

Figure 8.6. This comparison of the two objectives for calculating minimal internal fluxes shows the number of nonzero fluxes with respect to cultivation time.

The sum of flux magnitudes appears to be a good measure of overall metabolic activity. It seems to capture cell physiology much better than the minimal number of reactions. The associated number of reactions is less sensitive to the choice of time points (data not shown) and the time-traces appear smoother. If the metabolic burden is proportional to the number of enzymes, the proposed objective captures the relative difficulty of growth on different substrates as is indicated by the specific rate of biomass formation.

The fact that blue and white spheres touch at the end of growth on glucose indicates that there is some merit to the total enzyme restriction (Shlomi *et al.*, 2011). During the adaptation period, the organism is probably as conservative with new enzymes as possible, thereby sticking to the white spheres. Once the mRNAs have been produced, the system is back onto the blue track. At the end of growth on ethanol, the organism is more or less in the same state as before the glucose pulse except it probably had a little more non-essential enzymes tagging along.

8.4 Modeling fermentation in SSF

For the purposes of testing simultaneous saccharification and fermentation a small model of yeast growth on glucose was added to one of the hydrolysis models. Creative scaling was required to change the units from g/kg mixture to g/kg water.

Glucose uptake was assumed to be limiting bacterial growth and thus described by a Monod equation. Ethanol inhibition on growth was described in (Luong, 1985). Complete extinction by cell death was not necessary to include, only the reduced activity due to ethanol. Regretfully, this was enough to limit the maximum accumu-

lated concentration of ethanol to approximately 90 g/kg mixture regardless of the total amount of glucose provided.

The production of the primary carbon products (carbon dioxide, acetate, ethanol, and glycerol) and the minor acids (formic, succinic, malic, and lactic) were described using yield factors. This model contained the pH controller's base consumption as one of the dependent variables.

The SSF model predictions were very similar to the results presented in those of Song *et al.* (2009) and van Zyl *et al.* (2011).

8.5 Mathematical formulation of bioreactor model and parameter estimation

A mathematical framework for a fermenting bioreactor simulation, parameter estimation and optimization was further developed. The bioreactor model (DFBA model) is divided into three compartments: 1) macroscopic extracellular model describing mass, energy balances in reactor, 2) microscopic intracellular model describing metabolic fluxes inside cells and extracellular transfer fluxes, 3) models for flux boundaries and metabolite uptake rates. Model and compartments are illustrated in Figure 8.7.

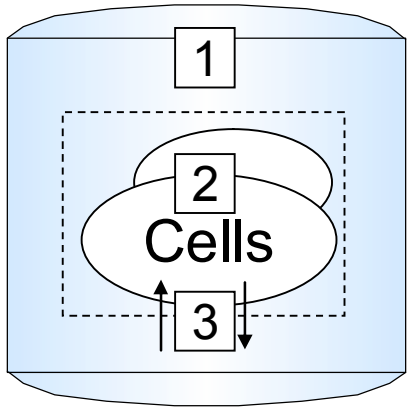

Figure 8.7. Bioreactor model compartments. 1) macroscopic extracellular model, 2) microscopic intracellular model, 3) models for flux boundaries and metabolite uptake rates.

The macroscopic model depicts a typical CSTR reactor with extensions for dynamic batch and fed-batch behavior. Microscopic model is a steady-state stoichiometric model of a micro-organism, where the metabolite fluxes inside the cells are calculated using flux balance analysis (FBA). FBA is based on reaction stoichiometries within the metabolic network, and some of the metabolic fluxes are constrained based on e.g. maximum metabolite uptake rates. The solution for all

the fluxes is found by finding a set of fluxes that maximize the cellular growth. The models for flux boundaries and metabolite uptake rates are usually based on empirical correlations. The overall model is formulated as a set of differential and algebraic equations (DAE).

An important step in constructing a predictive bioreactor model is estimating the model parameters from experimental data. The number of parameters in DFBA models is often rather small, say around 10, but the parameters may be strongly dependent on each other. Parameters are also often estimated from limited data. We have developed a gradient based parameter estimation method for DFBA models complemented with a parameter selection algorithm (Leppävuori *et al.,* 2011). These methods allow one to fit the model to target system and the model can be used to predict the behavior of the process. DFBA models have been used to e.g. design the sugar feeding strategy in fed-batch fermentation (Hjersted & Henson, 2006). The parameter estimation method is computationally efficient even with today's largest genome-scale metabolic networks, and thus the latest available knowledge of the cell metabolism can be utilized in fermentation process design. Example of measured and simulated data is presented in Figure 8.8.

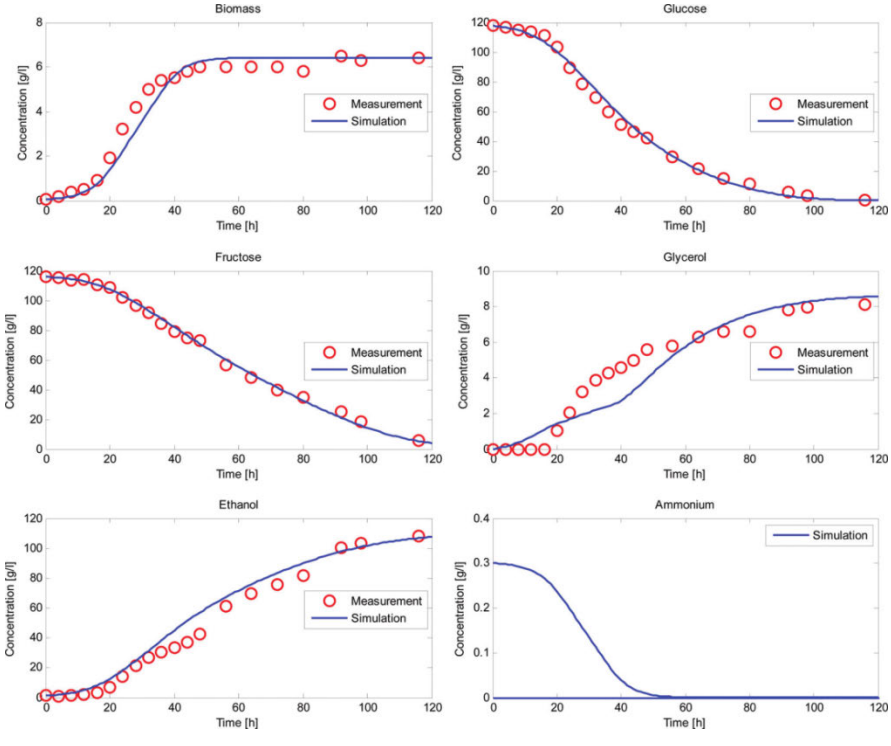

Figure 8.8. Example of simulated and measured metabolite concentrations in bioreactor (Leppävuori *et al.*, 2011)

8.6 Isobutanol as an example of novel fermentation product

The yield, titer, and production rate of novel products can be predicted based on dynamic flux balance analyses and experimental data (Tohmola *et al.,* 2011; Duarte *et al.,* 2004). Predictive physiological modeling is based on combining an extracellular process model (ordinary differential equations) with intracellular flux balance (linear programming) to accurately describe growth on various substrates, self-limiting factors, and product titer (Herrgard *et al.,* 2008). These tools were applied for comparing n-butanol to isobutanol and for predicting the realistic yields of ethanol and isobutanol.

8.6.1 Isobutanol versus n-butanol

There are many ways to evaluate the general feasibility of a novel path. One of the first things to do is to calculate the optimal theoretical yield. A net path stoichiometry is generated in the process. Altogether six paths were investigated in relation to this work. The paths link two substrates (glucose and xylose) with three products (isobutanol, n-butanol, and ethanol) as shown by the stoichiometry in Table 8.1.

Table 8.1. The net path stoichiometry for six paths spanning two substrates and three products.

Substrate	**glucose**	**glucose**	**glucose**	**xylose**	**xylose**	**xylose**
Product	**ethanol**	**n-butanol**	**isobutanol**	**ethanol**	**n-butanol**	**iso-butanol**
Glucose	-1	-1	-1			
Xylose				-1	-1	-1
H_2O			1			0.83
H^+		3.43			2.86	
CO_2	2	2.57	2	1.67	2.14	1.67
Ethanol	2			1.67		
n-Butanol		0.86			0.71	
Isobutanol			1			0.83

A path profile is a cofactor-unbalanced subset of the reactions involved in the net path stoichiometry. The path profiles utilizing glucose are given in Table 8.2.

Table 8.2. The path profiles for six pertinent subpaths from glucose to product.

Substrate	glucose	pyruvate	ATP	pyruvate	pyruvate	pyruvate
Product	**pyruvate**	**CO_2**	**ADP**	**ethanol**	**n-butanol**	**isobutanol**
Glucose	-1					
Pyruvate	2	-1		-2	-2	-2
H_2O	2	-2	-1		1	1
H^+	4	4		-4	-2	-2
CO_2		3		2	2	2
NAD	-2	-5		2	4	2
NADH	2	5		-2	-4	-2
ADP	-2	-1	1			
ATP	2	1	-1			
Pi	-2	-1	1			
Ethanol				2		
n-Butanol					1	
Isobutanol						1

The carbon efficiency of the path to n-butanol is clearly lower than for isobutanol or ethanol. This can be deduced from the row for carbon dioxide in Table 8.1. Furthermore, the path produces a large amount of protons as another byproduct. This massive generation rate of protons would quickly acidify the intracellular space and cause severe stress to the organism. The acidification of the intracellular space and the need for more reducing equivalents than glycolysis provides (evident by for NADH in Table 8.1) are the primary reasons why organisms also produce ethanol and acetone in the ABE-process. Reducing equivalents are carried by the Nicotinamide Adenine Dinucleotide cofactors denoted NADH for the reduced form and NAD for the oxidized form.

The regeneration of excess Adenosine Tri-Phosphate is here represented by the hydrolysis of ATP (third path in Table 8.2), but would in reality go to the upkeep of cellular structures. The upper limit of productivity, *i.e.* the maximum production rate of ethanol and isobutanol is determined by the ATP maintenance requirement at zero growth. Since the subpaths from pyruvate to all three product use NADH, they compete directly with oxidative phosphorylation. Oxidative phosphorylation is active whenever oxygen is available, making these processes most product-efficient at low oxygen (anaerobic) conditions.

8.6.2 Isobutanol predictions

The theoretical yield is never achieved in practice because the cell needs to repair itself and actively maintain the desired functionality by producing more enzymes to replace those that are spontaneously degraded. This implies a small growth requirement. Since large metabolic changes occur when the organism adapts from anaerobic conditions to microaerobic conditions, the average cell in industrial

anaerobic cultivations is maintained below this limit. The lowest predicted product yield, *i.e.* the most probable outcome, is at the very edge where anaerobic meets microaerobic.

The major oxygen-dependent growth stages are anaerobic, microaerobic, and aerobic. Computationally, anaerobic states are characterized by only using oxygen for such reactions that absolutely must have oxygen and whose products can not be produced by any other means. Fully aerobic is computationally equivalent to having all reducing equivalents balanced out by oxidative phosphorylation. Microaerobic is computationally the intermediate, where reducing equivalents are regenerated by both oxidative phosphorylation and *e.g.* alcoholic fermentation.

By varying the ratio of oxygen versus glucose, the three growth states can be computed by optimizing biomass yield as a primary objective and the product yield as a secondary objective. The realistic yields are then obtained as the yields at the bend in the curve between anaerobic and microaerobic states. The predicted realistic yields are listed in Table 8.3. The production path to isobutanol can be implemented either in the cytosol or in the mitochondrion. These two predictions differ because of the compartment-specific cofactor balances.

Table 8.3. Realistic yields predicted by flux balance analysis.

Substrate	Glucose	Glucose	Glucose	Xylose	Xylose	Xylose
Product	ethanol	isobutanol	isobutanol	ethanol	isobutanol	isobutanol
Compartment	cytosol	cytosol	mitochondria	cytosol	cytosol	mito-chondria
Glucose	-1	-1	-1			
Xylose				-1	-1	-1
H_2O	0.281	1.078	1.096	0.093	1.621	0.837
H^+	0.110	0.059	0.072	0.037	0.157	0.002
CO_2	1.751	1.867	1.836	1.584	1.775	1.661
Dry biomass	0.022	0.012	0.014	0.007	0.031	0.001
Ethanol	1.731			1.578		
Isobutanol		0.928	0.912		0.526	0.830

8.6.3 Maximum titer

The titer for isobutanol is limited by the toxicity of the product. This limitation has something to do with the membrane. Presumably, the membrane integrity is jeopardized when the amount of isobutanol in the membrane exceeds 20 mol-%. This value is roughly every fifth membrane surface molecule, indicating that each membrane lipid is adjacent to two foreign molecules. The relative reduction of charged membrane lipids lowers the screening potential that keeps ions from crossing the membrane. The amount of isobutanol in the membrane was predicted using the octanol-water partition coefficient. The octanol-water partition coefficients were estimated based on Hammett relations using commercial software [ACD Labs Structural Designer]. The molecular weights, the partition coefficients

and the maximum titers predicted using Equation (8.6.1) are tabulated in Table 8.4 for select products of interest.

$$c = \frac{x\,\rho_{\text{Octane}} MW}{(1-x)\,M_{\text{Octane}}\,10^{\log P}}$$
$$x = 0.2$$
$$\rho_{\text{Octane}} = 0.8\ g\,L^{-1}$$
$$M_{\text{Octane}} = 114\ g\,mol^{-1} \qquad (8.6.1)$$

Table 8.4. The molecular weight (MW), the base 10 logarithm of the octanol-water partition coefficient (LogP), and the predicted maximum titer for select products of interest.

Product	MW [g mol^{-1}]	LogP	Maximum titer [g (kg water)$^{-1}$]	Uncertainty interval
Ethanol	46	-0.1	102	80–184
Isobutanol	74	0.68	27	18–42
n-Butanol	74	0.84	19	12–29

Isobutanol titers reported in literature are 4.9 ± 0.1 g/L (Smith *et al.,* 2010), 2.62 g/L (Li *et al.,* 2011), and 10 ± 0.5 g/L (Baez *et al.,* 2011). The ethanol tolerance has been reported to be somewhere between 90 g/L and 140 g/L (Pereira *et al.,* 2011).

8.6.4 Thermodynamic analysis

The path from pyruvate to isobutanol was analyzed with energy diagrams. An energy diagram is a graphical visualization of Gibbs energy changes that are somehow connected to each other. Each bar in Figure 8.10 represents one predicted Gibbs energy change. The six reactions are described in Table (8.2). The lumped reactions R2R3 and R4R5 are the sum of the individual reactions R2, R3, R4, and R5. The Gibbs energy change for a reaction ($\Delta_r G$) is calculated as a sum of the Gibbs energies of formation ($\Delta_f G_k$) for reactants and products weighted by their stoichiometric factors (ν_k) in any given reaction.

$$\Delta_r G = \sum_k \nu_k \Delta_f G_k \qquad (8.6.2)$$

Five estimates for the transformed Gibbs energies of formation were obtained. The Gibbs energy of formation for the gas phase constituents were estimated with the methods by Marrero and Gani (MG), by Constantinou and Gani (CG), and by Joback and Reid (JR) as implemented by the ProPred software component within the ICAS package. The Gibbs energy of solvation was estimated using the COS-

MOtherm software. The Gibbs energy of formation for the aqueous constituent ($\Delta_f G_{k,aq}$) is the sum of the Gibbs energy of formation for the gaseous constituent ($\Delta_f G_{k,g}$) and the Gibbs energy change of solvation ($\Delta_s G_k$) according to the relation:

$$\Delta_f G_{k,aq} = \Delta_f G_{k,g} + \Delta_s G_k \tag{8.6.3}$$

Conventional Gibbs energies are reported on a particular scale where the standard reference energy of the hydrogen ion is zero at a concentration of 1 M. The convention in biochemistry is to use an energy scale where the energy of the hydrogen ion is zero at a concentration corresponding to pH 7. Energies on the former scale can be translated to energies in the latter scale by the use of a Legendre transform, where the energy of the hydrogen ion is determined from the pH and total number of hydrogen component (n_H) in the reaction system:

$$\Delta_f G'_{k,aq} = \Delta_f G_{k,aq} - n_H \Delta_f G_{H,aq} \tag{8.6.4}$$

$$\Delta_f G_{H,aq} = \Delta_f G^{\circ}_{H,aq} + RT \ln\left(10^{-pH}\right) \tag{8.6.5}$$

The transformed Gibbs energies of formation have been directly estimated by a group contribution method developed by Jankowski *et al.* (2008). The method is largely based on the apparent Gibbs energy changes and apparent equilibrium constants found in the Standard Reference Database 74, compiled by the National Institute of Standards and Technology (Beckham *et al.,* 2011). A recent group contribution approach made from the same data and estimated acidity constants (pK_a values) will estimate traditional Gibbs energies of formation (Noor, 2011, personal communication).

The aqueous Gibbs energies were also estimated via the solid state by first estimating the Gibbs energy of formation for the solid substances and then adding estimates for solubility. Since the solid state prediction method was not directly applicable to these compounds, the two alternate ways were chosen for each *ad hoc* extrapolation alternative given. In the end, this did not have that big of an effect on the overall predictions.

The pathway investigated here is the path from pyruvate to isobutanol. The path consists of six enzymatically catalyzed reactions, of which two are catalyzed by the same enzyme. The seven metabolites are shown in Figure 8.9. Reactions 3 and 6 use NADH, while reactions 1 and 5 produce one carbon dioxide molecule. Reaction 4 produces one water molecule. Every isobutanol molecule requires two pyruvate molecules. Table 8.1 contains the overall stoichiometry of this path.

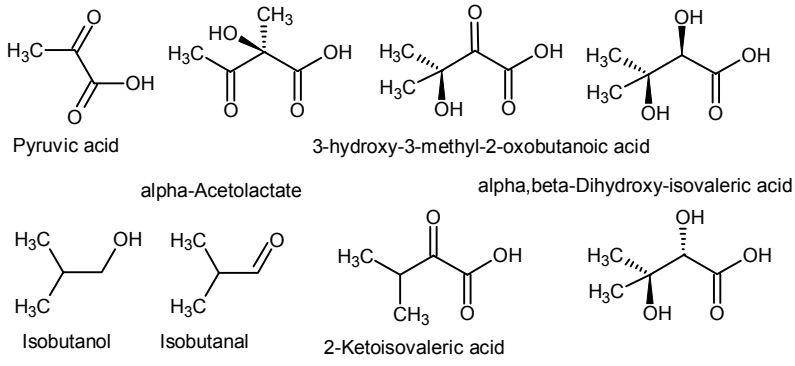

Figure 8.9. Molecular structures of the seven metabolites in the path from pyruvate to isobutanol. The metabolite alpha,beta-dihydroxy-isovaleric acid has two plausible stereochemical structures. The top row represents the three first reactions and the second row the three last reactions.

The Gibbs energy changes for reactions in the path from pyruvate to isobutanol are calculated from the five estimates of the Gibbs energies of formation and displayed in the energy diagram in Figure 8.10. The reactions are denoted R1 thru R6. The notation R2R3 implies the combination of reactions 2 and 3, *i.e.* skipping an intermediate. These lumped reactions are particularly useful to use when the energy of an intermediate could not be generated.

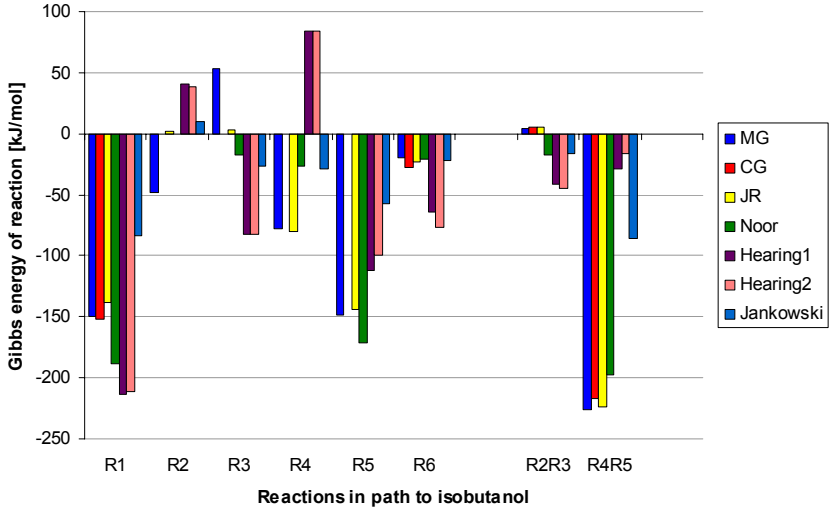

Figure 8.10. An energy diagram of the reactions participating in the path from pyruvate to isobutanol. The path has 6 steps, but since thermodynamic properties could not be generated for all intermediates, two lumped reactions are also included. See text for explanations.

As for most biochemical paths, the least thermodynamically feasible reactions occur in the beginning of the paths where carbon is committed to the path. Although the MG-method overestimates the stability of 3-hydroxy-3-methyl-2-oxobutanoic acid, the predictions for the combined reaction R2R3 are surprisingly similar for all six methods.

As customary to biochemical paths, the commitment to this pathway in the beginning of it is the thermodynamic bottleneck. The high sensitivity to substrate, enzyme, and cofactor concentrations enable optimal control of the flux thru the pathway.

This path is active at nearly any conditions (pH, NADH/NAD cofactor ratio, total concentration). Large amounts of NAD will cause the path to slow down.

8.6.5 Discussion

While there is a combinatorial number of ways to use the property predictions by ICAS, ACD Labs, and COSMOTHERM, the method above was chosen for having the least amount of extra assumptions of ideality or otherwise. The spread in predictions is large for any given metabolite, but smaller for reactions and Gibbs energy changes in general.

Although a single condition was use to compare the selected prediction methods, it represents well all conditions considered. The influence of metabolite-specific concentrations has a marginal effect on all but the rate-controlling enzyme. It is concluded that the rate of catalysis is a bigger concern than the feasibility.

8.6.6 Summary

Based on the optimal theoretical yield and the stoichiometry of the net path, isobutanol is clearly the better product compared to n-butanol. The production of isobutanol has the greatest potential in an environment with little oxygen. The maximum titer for isobutanol was predicted to be around 27 grams per kilogram water. The path from pyruvate to isobutanol is thermodynamically feasible under a wide range of conditions. A single enzyme controls the commitment of carbon flux to this pathway and it therefore important for studying the regulation of this path.

9. Bark-based bioethanol production

The process under study was a bark biorefinery depicted in Figure 9.1. It consists of hot water extraction of bark at ambient pressure to isolate the bark speciality compounds for use in adhesives or fine chemicals. The extract, containing polyphenols, stilbenes and mono- and polysaccharides, goes through solids separation followed by ultrafiltration and stilbene separation. The polyphenols fraction is then evaporated and finally spray-dried. The produced crude tannin contains ca. 25% polysaccharides, mostly glucose, the rest being predominantly tannin and tannin-like polyphenols. The extracted and pressed bark is diluted using permeate from tannin extract ultrafiltration, neutralized using lime, cooled to 45°C and sent to enzymatic hydrolysis. The hydrolysed bark suspension is fermented and the produced alcohol concentrated using distillation. The distillation bottoms are concentrated to 30% dry content using multiple effect evaporation that is heat integrated with the distillation column and sent to combustion.

The extraction of bark is done in a counter-current fashion. The bulk of fresh water entering the process is used to dilute and wash the extracted bark before and during pressing. The filtrate from pressing is recycled back to extraction.

The main design variables are:

1. Fresh water intake. It has an influence on the extraction consistency, the concentration of dissolved bark component in the crude extract and the amount of waste water produced.
2. Hydrolysis consistency. It has an influence on the sugar concentration after hydrolysis and hydrolysis yield. The influence on yield is not considered in the present model.
3. Steam efficiency (kg steam / kg water evaporated) in various evaporation tasks.
4. Closing of water circulations by evaporating surplus UF-filtrate from tannin and stilbene recovery. The closing of water circuits eliminates the cost associated with waste water treatment and somewhat increases the ethanol yield. The decision variables 1 to 3 are continuous, whereas variable 4 is a binary one. There is no point of building an evaporation plant to partially close the water circuits.

In addition to the above design variables there are a number of uncertain process parameters like yields, separation efficiencies, dry contents after solid-liquid separations, etc. A third set of parameters influencing the profit function only is composed of the prices of utilities consumed and those of products produced.

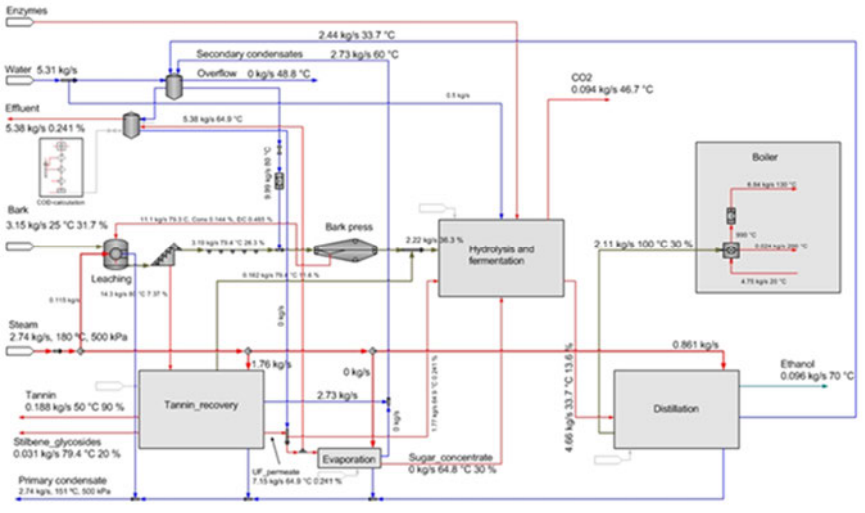

Figure 9.1. The Bark Biorefinery.

As described in Chapter 3.2 the bark biorefinery process was studied with two methods of global sensitivity analysis, namely the method of separating model parameters and variables and the Elementary Effects (EE) method. The Elementary Effects sensitivity analysis was conducted by generating 204 samples (each consisting of values for the 33 factors, see Appendix E), simulating all the samples and performing post-simulation statistical analyses. In the EE method the post-simulation analysis comprises of calculating for each factor in two numerical indicators, namely μ_i and σ_i. These two values are used to determine whether the parameter is important or negligible. Typically a so-called (μ,σ)-plot is used, in which each parameter is plotted using its μ_i as the x-coordinate and σ_i as the y-coordinate. A typical (μ,σ) plot is shown in Figure 9.2. An example of Elementary Effects (μ,σ)-plot. The figure also shows typical interpretations of the results. In the plot of Figure 9.3 also the codes are shown for the five most important factors.

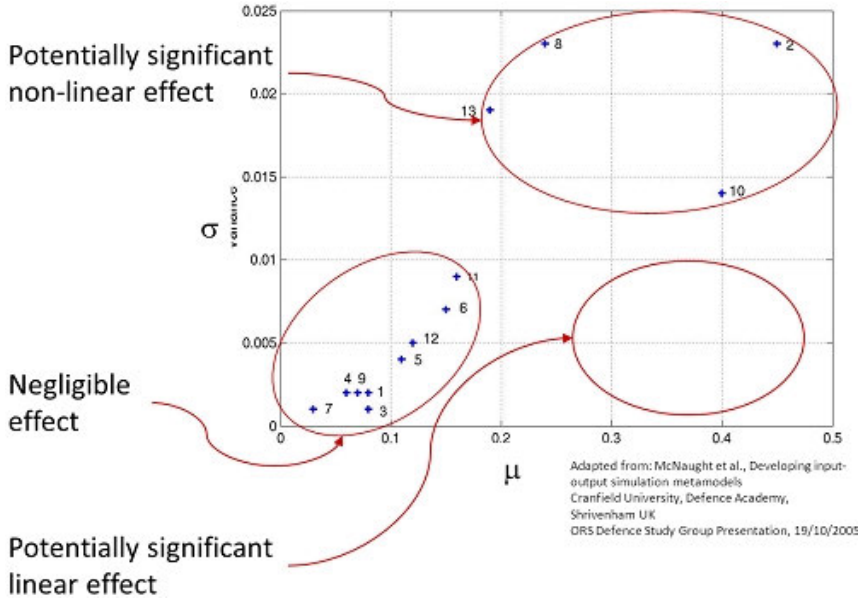

Figure 9.2. An example of Elementary Effects (μ,σ)-plot.

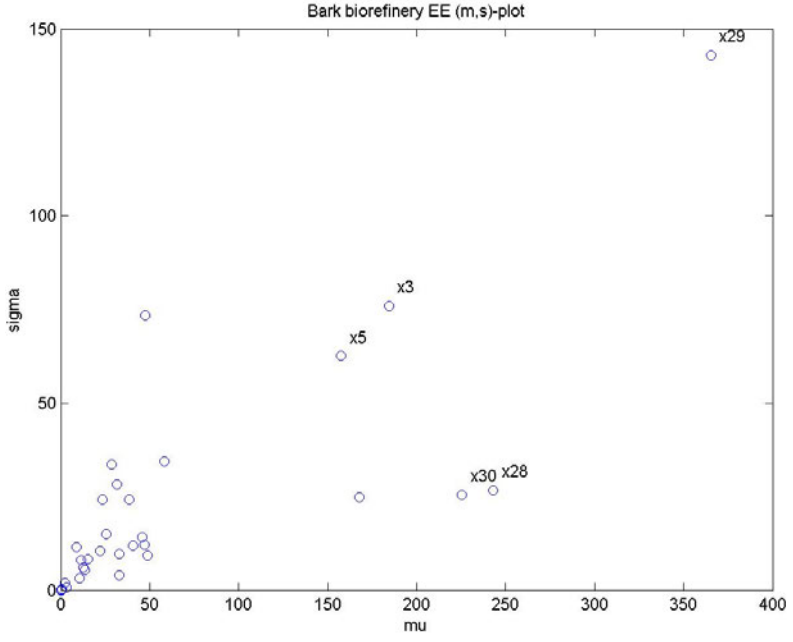

Figure 9.3. Bark biorefinery EE (μ,σ)-plot with names of five most important factors as explained in the text.

The conclusion is that factor coded by x_{29} is by far the most important one. This parameter is the steam unit price (€/ton). The next four factors (x_3, x_{28}, x_{30} and x_5) are the fresh water feed to leaching, fuel price, tannin price and steam consumption in the evaporation.

The conclusion of the Elementary Effects analysis is well in line with the previous bark biorefinery sensitivity analyses.

10. The comparison of greenhouse gas emissions from barley straw based ethanol and butanol

The aim of this study was to assess the difference between the greenhouse gas (GHG) emissions of ethanol and butanol by interpreting the RED methodology. First, the greenhouse gas emissions of both bioliquids were calculated separately. Then the difference between the GHG emissions of ethanol and butanol was deduced. The values for the necessary parameters were gathered from literature sources. When no reliable data was found, the parameters were assumed roughly based on expert guesses. The default values for process parameters of biofuel processing are derived from in-house data of VTT. The uncertainty analysis was carried out using a stochastic Monte Carlo simulation. The results for the GHG emissions of ethanol and butanol and the difference between them are presented as probability distributions.

10.1 Concept description

The main features of ethanol and butanol processes are outlined in Figure 10.1. The raw material for these bioliquids was assumed to be barley straw. First, the barley straw is milled to small pieces and then the milled straw is blended with water. Pretreatment is necessary to allow the enzymes to hydrolyse the cellulose to monometric sugars. The selected pretreatment method is steam explosion with addition of acid catalyst H2SO4 (sulphuric acid). Then the raw material flow is diluted with water and the pH is adjusted with NaOH (sodium hydroxide). Enzymatic hydrolysis converts the hemicelluloses and the celluloses to monometric sugars which are then fermented by yeast. In the process the solid residue (co-product 1) is filtrated from the product flow before ethanol or butanol is distilled. The residue from distillation (co-product 2) is concentrated in the evaporation plant. The ethanol/butanol process is integrated to combined heat and power (CHP) plant.

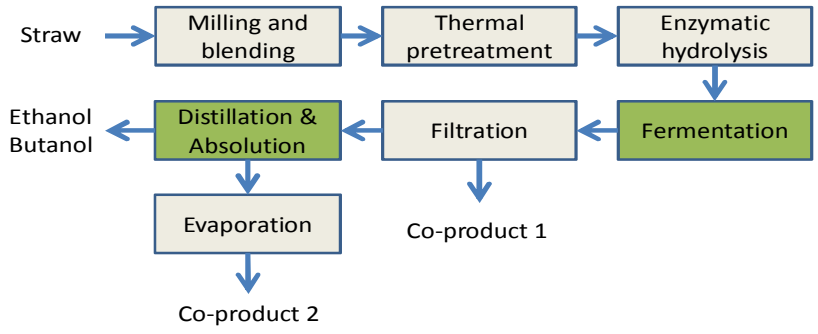

Figure 10.1. Simplified block diagram of ethanol and butanol processes. Fermentation and distillation unit processes (marked in green) are different in ethanol and butanol processes while the rest are identical.

The process parameters for ethanol and butanol processes are presented in Table 10.1. The plant capacity is 160 000 tons of straw in dry matter per year. The operation rate of the plant was assumed to be 95%.

Table 10.1. Process parameters for ethanol and butanol processes.

Name	Unit	Min. value	Default value	Max. Value	Distribution
Utilization rate of the plant	%	90%	95%	99%	Beta
Electricity consumption of milling	kW	203	290	377	Normal
Yield of milled straw	%	98%	99%	100%	Beta
Consumption of high pressure steam (200°C) in pretreatment	MW	10	15	19	Normal
Consumption of H_2SO_4	$g/kg_{d.m}$	1,75	2,50	3,25	Normal
Consumption of NaOH	$g/kg_{d.m}$	20,51	29,30	38,09	Normal
Electricity consumption of pretreatment	kW	21	30	39	Normal
Consumption of enzymes	$g/kg_{d.m}$	136	194	252	Normal
Electricity consumption of hydrolysis	kW	21	30	39	Normal
Electricity consumption of fermentation	kW	581	830	1079	Normal
Low pressure steam (85°C) from pretreatment to fermentation	MW	2,0	2,9	3,8	Normal
Consumption of low pressure steam (85°C) in fermentation	MW	1,8	2,5	3,3	Normal
Consumption of yeast	$g/kg_{d.m}$	211	301	391	Normal
Bioliquid loss included in co-product 1	$g/kg_{d.m}$	5	10	20	Weibull
Yield of co-product 1	$g_{d.m}/kg_{d.m}$	358	512	666	Normal
Yield of dilute ethanol	$g_{d.m}/kg_{d.m}$	139	199	258	Normal
Yield of dilute butanol	$g_{d.m}/kg_{d.m}$	119	170	221	Normal
Consumption of middle pressure steam (140°C) in distillation of ethanol	MW	7	10	13	Normal
Consumption of middle pressure steam	MW	12,6	18	23,4	Normal

Name	Unit	Min. value	Default value	Max. Value	Distribution
(140°C) in distillation of butanol					
Yield of bioliquid in distillation	%	92%	95%	98%	Beta
Yield of co-product 2	$g_{d.m}/kg_{d.m}$	213	304	395	Normal
Low pressure steam (100°C) from pretreatment to evaporation	MW	8,1	11,5	15,0	Normal
Consumption of low pressure steam (100°C) in evaporation	MW	7,2	10,3	13,4	Normal
Electricity consumption of evaporation	kW	63	90	117	Normal
Yield of co-product 2 from evaporation	%	96%	97%	98%	Beta

10.2 Parameters for GHG emission calculations

10.2.1 Emissions from the cultivation of raw materials

According to the RED, emissions from the cultivation of raw materials shall include emissions from the cultivation process itself; from the collection of raw materials; from waste and leakages; and from the production of chemicals or products used in the cultivation.

It has been estimated that collection of straw diminishes the soil nitrous oxide emissions. For example Malhi *et al.* (2006), Malhi & Lemke (2007) and Gregorich *et al.* (2005) have studied these emissions. The N_2O emissions depends e.g. on precipitation, fertilization and soil type. In this study the figures used for N_2O loss per hectare are based on Malhi & Lemke (2007) because no country-specific data for Finland was available. At the same time the carbon balance of the field is estimated to reduce due to straw collection. The changes in soil carbon stocks due to raw material cultivation or harvesting have not been considered in the default values of the RED. Neither does RED give instructions on how to calculate these emissions or which time frame to use in calculations. Despite the lack of instructions, these emissions were taken into account in this study, because they are closely related to the emissions from the cultivation of raw materials. In this study it was estimated that after 20 years 19% of the carbon from straw is still remaining in the soil if the straw is left on soil. This figure is based on Swedish COUP-model (Jansson & Karlberg 2001; Salo, 2007).

The loss of nutrients in the soil due to collecting of straw is compensated by fertilization. The amount of fertilizer needed to compensate the loss was estimated using figures presented in Malhi and Lemke (2007). The emissions from the manufacturing of fertilizers are based on Mäkinen *et al.* (2006). The emissions from production and transportation of the fertilizer, and the direct and indirect N_2O emissions from soil were taken into account (Mäkinen *et al.,* 2006).

The emissions from straw collecting and transportation, from fertilization, from changes in soil carbon stock and emission saving from diminished N_2O emissions

are presented in Table 10.2. Emissions from straw collection and transportation are based on Mäkinen *et al.* (2006). The emissions from diesel fuel production, distribution and dispensing are based on Edwards *et al.* (2007). The CO_2 emissions from diesel fuel use were calculated using emissions coefficient of 68 gCO_2/MJ (Statistics Finland, 2007). CH_4 and N_2O emissions estimates were 0.16 and 1.12 g/MJ respectively. The lower heating value used for diesel was 42.9 MJ/kg and the density 830 g/l.

Table 10.2. Parameters for the emission calculation from cultivation of raw materials.

Name	Unit	Min. value	Default value	Max. Value	Distri-bution
N_2O emission saving due to straw collecting	g_N/ha	-50	25	143	Weibull
Collecting rate of straw	%	60%	70%	80%	Normal
Yield of straw	kg_{wet}/ha	1450	3603	6100	Normal
Yield of straw per hectare	kg/ha	1041	1225	1409	Normal
Moisture content of straw	%	12%	13%	14%	Beta
Heating value of straw as received	MJ/kg	16,7	17,1	17,4	Normal
Increased amount of fertilization due to straw collecting	kg_N/ha	2	14	27	Weibull
Carbon content of straw	%	45%	46%	47%	Beta
Carbon remaining in soil after 20 years if straw isn't collected	%	0%	10%	19%	Beta
CO_2 emissions from fertilizer production	$kgCO_2$/t	455	506	557	Normal
N_2O emissions from fertilizer production	kgN_2O/t	1,5	2,0	3	Normal
CH_4 emissions from fertilizer production	$kgCH_4$/t	0,1	0,2	0,2	Normal
Diesel consumption in transportation of fertilizers	g/t, km	18,5	22,4	29,9	Weibull
Diesel consumption in collecting of straw	l/MWh	0,6	0,8	0,9	Normal
Diesel consumption in transportation of straw	l/MWh	0,5	1,0	2,0	Weibull
Diesel consumption in breaking of bundles	l/MWh	0,9	1,0	1,1	Normal
N_2O emissions from soil	gN_2O/kg_N	4	26	112	Weibull
Emissions from production of diesel fuel	tCO_{2eq}/l	2953	3281	3609	Normal

10.2.2 Emissions from processing

In the production of ethanol and butanol, emissions from processing stage consist of electricity and steam generation, and manufacturing of yeast, enzymes, sulphuric acid and sodium hydroxide. Parameters for these emissions are presented in Table 10.3.

In case 1 the electricity is supplied from outside to the bioliquid production plant, whereas in case 2 it is assumed to be produced within the production plant. The RED defines that in accounting for the consumption of electricity not produced within the fuel production plant, the greenhouse gas emission intensity of the production and distribution of that electricity shall be assumed to be equal to the av-

erage emission intensity of the production and distribution of electricity in a defined region (EC, 2009). The annual production-based CO_2 emission intensities of electricity consumption between 2000 and 2008 in the Nordic countries (lower limit), Finland (default value) and EU-27 (upper limit) were used (Soimakallio & Saikku 2012). The emissions of sulphur acid and sodium hydroxide are based on Mäkinen *et al.* (2006). The emissions from enzyme production are based on Ecoinvent (2010) and emissions from yeast production are based on Nielsen *et al.* (2007).

In both cases the steam is provided by CHP plant combusting the two ligneous co-products. The need for start and back-up fuel was estimated to be 1% of the energy content of the co-products. Heavy fuel oil was assumed to be start and back-up fuel.

Table 10.3. Emissions from chemical inputs to the process.

Name	Unit	Min. value	Default value	Max. Value	Distribution
Emissions from H_2SO_4	$kgCO_2/kg$	137	195	254	Normal
Emissions from NaOH	$kgCO_2/kg$	836	1195	1553	Normal
Emissions from yeast production	$kgCO_{2eq}/kg$	0,51	1,01	1,52	Normal
Emissions from enzyme production	$kgCO_{2eq}/kg$	1	6	10	Normal

10.2.3 Emissions from transportation and distribution

According to Edwards *et al.* (2007) emissions from diesel fuel transportation and distribution equals 1.1 gCO_{2eq}/MJ. This figure includes emissions from transport (barge, rail, and pipeline), depot, distribution and dispensing. It was assumed that there is no difference in corresponding emissions of ethanol and butanol compared to diesel fuel. These emissions were assumed to be normally distributed within the uncertainty range of ±30%. The heating values for ethanol and butanol are 27 MJ/kg and 33 MJ/kg, respectively.

10.2.4 Emission saving from excess electricity from cogeneration

The emission saving from excess electricity from cogeneration shall be taken into account when the fuel used for cogeneration is an agricultural crop residue (EC, 2009). Straw is an agricultural residue but in both cases the CHP plant uses co-products 1 and 2 for cogeneration which actually are not straw but ligneous co-products of the process. Thus it is not clear whether these ligneous co-products should be handled as agricultural crop residue. The emission saving from excess electricity cogeneration was assumed to be taken into account.

According to the RED when calculating the excess electricity, the size of the cogeneration unit should be assumed to be the minimum necessary for the co-generation plant to supply the heat that is needed in processing the fuel. The greenhouse gas emission saving should be taken to be equal to the amount of

greenhouse gases that would emitted when the equal amount of electricity is produced in a power plant using the same fuel as the cogeneration plant. (EC, 2009.)

The emission saving was calculated assuming that straw is used in power plant to produce the same amount of electricity than the excess electricity from cogeneration. The efficiency for electricity production is based on the harmonised efficiency reference value for separate production of electricity from agricultural biomass in a production plant that is constructed between 2006 and 2011 (COM 2007). The CH_4 and N_2O emission factors are based on Tsupari *et al. (*2007). The emission saving consists of emissions from straw collecting and transportation, fertilization, and changes in soil carbon stocks and the loss of N_2O emissions due to straw collection. The rest of the parameters needed in the calculation of the emission saving are presented in Table 10.4 (most of the parameters are already presented in Chapter 10.2.1 and they are not represented here).

Table 10.4. Parameters for the calculation of the emission saving from excess energy.

Name	Unit	Min. value	Default value	Max. Value	Distribution
Emissions from electricity production	gCO_2/kWh	96	221	449	Uniform
Consumption of HFO	%	0,5%	1%	4%	Beta
Primary energy need in HFO production	kJ_{prim}/MJ	979	1088	1197	Normal
Emissions from HFO production	gCO_{2eq}/MJ	77	86	95	Normal
Boiler efficiency	%	0,80	0,88	0,96	Normal
Electricity efficiency in CHP plant	%	35%	40%	45%	Beta
Electricity efficiency in condense power plant	%	20%	23%	25%	Beta
Heating value of co-products in dry matter	MJ/kg	19	21	22	Normal
N_2O emissions from combustion of co-products	mgN_2O/MJ	0	5	10	Normal
N_2O emissions from combustion of straw	mg/MJ_{fuel}	1	2	6	Weibull
CH_4 emissions from combustion of straw	mg/MJ_{fuel}	0,5	1,0	3,2	Weibull

10.3 Results

The results are presented as probability distributions. In case 1 (Figure 10.2), where the production plant and the CHP plant are considered as independent units, the emissions of butanol are slightly higher than emissions of ethanol in corresponding process. The small difference is due to a smaller yield of butanol and a greater energy demand in distillation of butanol compared with distillation of ethanol. The difference between the emissions of butanol and ethanol as per cents in cases 1 and 2 are presented in Figure 10.4

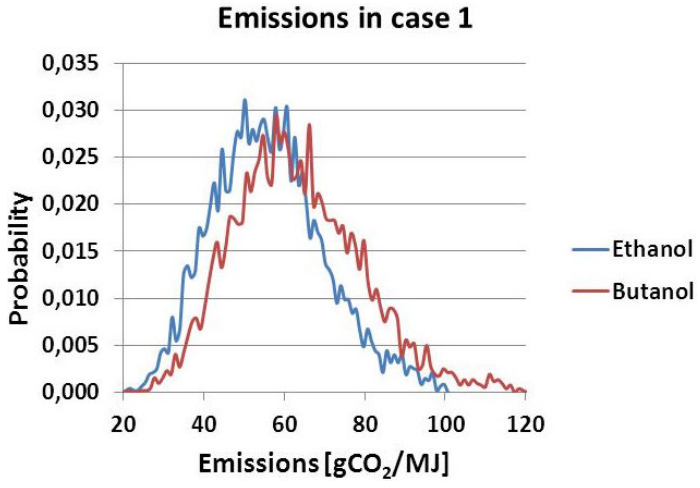

Figure 10.2. Emissions from ethanol and butanol production in case 1 where the production plant and the CHP plant is considered as independent units.

In case 2 (Figure 10.3), where the production plant and the CHP plant are considered as one combined unit, the emissions of butanol are significantly higher than the emissions of ethanol. The greater difference is mainly due to a system boundary setting. The butanol process consumes more steam than the ethanol process. Due to higher steam consumption, the amount of electricity sold to the grid is smaller in butanol case than in ethanol case. This leads to different results.

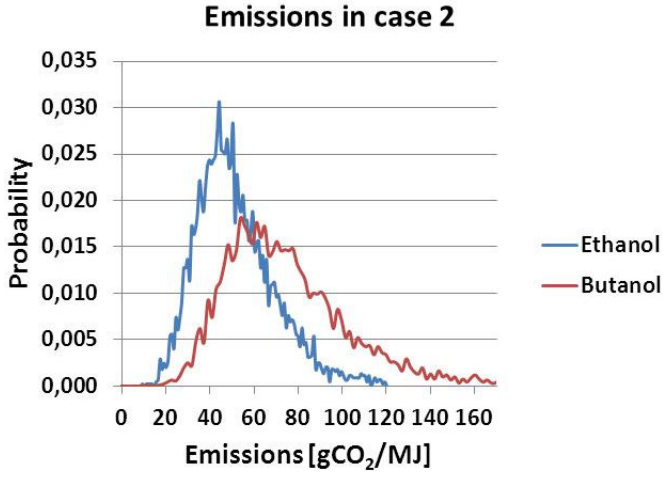

Figure 10.3. Emissions of ethanol and butanol in case 2 where the production plant and the CHP plant are considered as one combined unit.

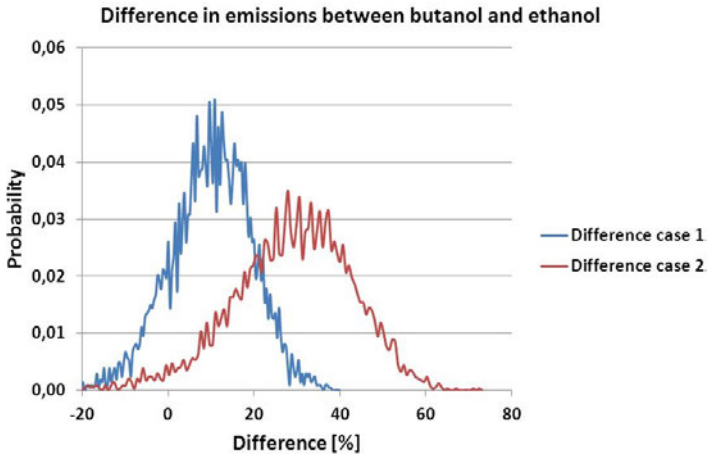

Figure 10.4. The difference of emissions from butanol production compared with ethanol production in case 1 and 2. Positive values indicate that the emissions of butanol are higher compared to those of ethanol.

The contribution of each parameter to the result was measured using Spearman's rank correlation. Only the correlations that have absolute value more or equal to 0.1 are presented in Table 10.5. In case 1 for both butanol and ethanol processes the most significant parameters are the amount of carbon remaining in the soil, the increased amount of fertilization, N_2O emissions from soil, and emissions from enzyme production. In case 2 the most significant parameters are yield of co-products, yield of ethanol/butanol and, emissions from enzyme production. The most significant parameters contributing the difference in emissions between ethanol and butanol are the yield of ethanol/butanol and the consumption of middle pressure steam in distillation.

Table 10.5. Spearman's rank correlations for each case.

Name	Ethanol case 1	Butanol case 1	Ethanol case 2	Butanol case 2	Difference case 1	Difference case 2
Carbon remaining in soil after 20 years if straw isn't collected	0,573	0,563	0,279	0,301		
Increased amount of fertilization due to straw collecting	0,354	0,355	0,161	0,182		
N2O emissions from soil	0,343	0,346	0,16	0,186		
Emissions from enzyme production	0,33	0,338	0,35	0,305		
Yield of dilute ethanol	-0,277		-0,307		0,708	0,537
Yield of co-product 2	-0,246	-0,27	-0,327	-0,326		
Yield of co-product 1	-0,126	-0,13	-0,479	-0,468		
Heating value of co-products in dry matter			-0,228	-0,215		
Consumption of high pressure steam (200°C) in pretreatment			0,203	0,191		
Boiler efficiency			-0,161	-0,182		
Consumption of low pressure steam (100°C) in evaporation			0,155	0,137		
Electricity efficiency in CHP plant			-0,138			
Consumption of middle pressure steam (140°C) in distillation of ethanol			0,1			-0,239
Consumption of middle pressure steam (140°C) in distillation of butanol				0,233	0,1	0,426
Yield of dilute butanol		-0,274		-0,322	-0,656	-0,579

10.4 Summary and inference

The renewable energy directive (RED) of the EU sets mandatory targets to increase the use of renewable energy by 2020 in the EU Member States. The RED introduced sustainability criteria for transportation biofuels and other bioliquids in order to ensure that their increasing production does not cause serious environmental and social problems. As a part of these criteria, the first ever mandated methodology was introduced to calculate GHG emission reductions resulting from the use of biofuels instead of fossil fuels. The main objective of this paper was to calculate the difference in GHG emissions of ethanol and butanol in corresponding process by applying the methodology provided in the RED. The bioliquid production was assumed to be integrated into a combined heat and power production (CHP) plant. The sensitivities and uncertainties were studied by setting various system boundaries and using stochastic modelling to explore the parameter uncertainty.

The emissions of butanol were overall higher than the emissions of ethanol. The difference is mainly due to a lower energy yield and higher consumption of

steam in distillation of butanol compared with ethanol. 20 year time frame was used for changes in soil carbon balances in every studied case. If the time frame is expanded, for example to 100 years, the emissions of ethanol and butanol would decrease significantly. Consideration of the bioliquid process and the CHP plant as one combined unit increases the emissions of ethanol and butanol. This is due to the fact that more emissions from CHP plant are attributed to the butanol/ethanol than when the units are considered separately. Though the setting of system boundary affects the results, the emissions of butanol seem to be higher than the emissions of ethanol despite the chosen calculation method.

The RED leaves room for setting various system boundaries and parameters. These may have significant impact on the GHG performance of biofuel or bioliquid. It is possible that two different biofuel chains with no significant differences in their characteristics give totally different emission results. In order to avoid this kind of situation more specific guidelines on setting the system boundaries and selecting the parameters for the calculation of actual GHG performance values of biofuels or bioliquids should be published to support the RED.

Liable actors may under certain conditions use the default values for GHG emissions and emission saving for the biofuels and other bioliquids provided in the RED. The default greenhouse gas emission saving of wheat straw ethanol is 85% (EC, 2009). The production of wheat straw ethanol can be seen as a very similar bioliquid process as barley straw ethanol considered in this study. The default value for emission reduction given in the RED is significantly higher than the results of this study, due to different assumptions. For example, changes in soil carbon stocks due to raw material cultivation or harvesting have not been considered in the default values of the RED. Thus, it could be justified to ignore them also when calculating actual GHG emissions. However, the treatment of changes in soil carbon balances influences the results. The appropriate time frame and method to account for changes in soil carbon balances should be fixed in the framework of the RED.

The difference between the emissions of ethanol and butanol is highly dependent on the yield of bioliquids and the energy demand in distillation. More accurate process data is needed in order to determine which bioliquid process produces less greenhouse gas emissions. In addition, not much emission data was available for enzymes and yeast. Especially, the emissions from enzyme production have relatively high contribution to the results. Therefore, the data for emissions from enzyme and yeast production should be more accurate.

11. Conclusions

The Biorefinery will eventually be a production plant consisting of separate unit operations and processes. For their modelling, not only the structure and dimensions of the various devices must be known, but it is essential to know the mecessary material properties as well as chemical reaction rate parameters of the bio-based process streams. Some of the less well known biorefining unit processes were studied in terms of these properties and parameters. The chosen unit processes were flash condensation of the fast biomass pyrolysis, thermal and enzyme catalysed hydrolysis of lignocellulosic biomass as well as fermentation of the forest biomass hydrolysis product. In addition, a flowsheet based mass and energy balance was developed for the bark biorefinery, the key factors of which were then assessed by using the Elementary efficiency sensitivity analysis. With the results from hydrolysis and fermentation models, a comparison of the greenhouse gas emissions from barley straw based ethanol and butanol was eventually made.

The biorefining processes appear generally similar to those found in petrochemical and biochemical industries; yet they include often a great number of molecules and their mixtures and many times as more or less dilute aqueous solutions. The reaction rates are often affected by mixture complexity or enzyme saturation. In general, however, dilute solutions and complex mixtures make down stream processing more determining for overall profitability than the reaction rates of bioprocesses.

In cases where the molecules as well as reactive processes and pathways can be identified, it is possible to use advanced estimation methods for their physicochemical properties. However, such complex mixtures as pyrolysis condensates it appears to be quite tedious to distinguish the salient components, which then could be utilised as representative for the modelling of desired unit operations. It remains a further challenge to develop reliable models to these mixtures to be utilised for e.g. the subsequent upgrading or fractionation processes.

Metamodelling could be used to support some of the necessary sensitivity analysis of the numerous model parameters while preparing for optimisation tasks; yet it was concluded that the phenomenon-based physicochemical modelling of the biorefining processes most often requires nonlinear mathematics and discontinuities e.g. due to phase changes. Such phenomena are difficult to include in the present metamodelling techniques and their modelling is in general well taken care of in chemical engineering methods.

With reasonable assumptions, flowsheet energy and mass balance modelling of the biorefinery becomes possible, sometimes also by using (unphysical) correlation or metamodels for those stages, to which a reliable phenomenon-based physical model is not accessible. Due to the great number of parameters, a sensitivity analysis is needed to support right conclusions. The model results are best applied for life-cycle and emission impact assessments.

12. Report summary

Biorefining unit processes were studied in terms of their physicochemical modelling properties and parameters. The material properties were estimated by using available equilibrium data for selected model compounds. The assessment was made by using the programs Aspen, ICAS, Cosmo-RS and Cosmo-SAC. The properties and parameters were further used in appropriate equation of state (EOS) models for the flash condensation of the fast biomass pyrolysis, thermal and enzyme catalysed hydrolysis of lignocellulosic biomass and fermentation of the hydrolysis product to bioethanol and biobutanol. Reaction kinetic modelling was applied for a number of unit processes including hydrodeoxygenation of the pyrolysis condensate, hydrothermal processing of lignocellulosisc biomass as well as for its low-temperature hydrolysis and fermentation of the hydrolysis solution. Optimisation of the fermentation model parameters was performed in co-operation with Carnegie-Mellon University.

Surrogate metamodelling could be used to support the necessary sensitivity analysis of model parameters while preparing for optimisation tasks; yet it was concluded that the phenomenon-based physicochemical modelling of the biorefining processes most often requires nonlinear mathematics and discontinuities e.g. due to phase changes, which are difficult to include in the present metamodelling techniques.

A mass and energy balance based cost estimation model was constructed for the biorefinery using bark as feedstock. A comparison of CO_2 emission from (straw-based) bioethanol and biobutanol production was made on the basis of the aforementioned fermentation models. The bark-based biorefinery producing mainly tannin and ethanol proved economically challenging. The results also show that biobutanol is in general more demanding (in terms of greenhouse gas emissions) than production of bioethanol. The viability of the chemical engineering simulation methodology in modelling the unit operations and processes of the biorefinery could be assessed. More work, however, is needed to gain a reliable material property base for such simulations, as well as for appropriate recognition of the dominant modelling parameters.

Acknowledgements

- Neste Oil Oyj
- Metso Power Oy
- Pöyry Finland Oy
- Vapo Oy
- Technology and Innovation Agency of Finland – Tekes

References

Abovsky, V. & Watanasiri, S. 1998. Mixing rules for van der waals-type equations of state based on activity-coefficient models. International Journal of Thermophysics 19(5), 1429–1446.

Abovsky, V. & Watanasiri, S. 1999. Equation of state mixing rule based on activity coefficient model: Explicit solution for 'finite pressure approach'. Fluid Phase Equilibria 158–160, 259–269.

Accelrys, I. 2007. From http://accelrys.com/products/datasheets/dmol3.pdf.

Ahlers, J. & Gmehling, J. 2001. Development of an universal group contribution equation of state: I. prediction of liquid densities for pure compounds with a volume translated Peng–Robinson equation of state. Fluid Phase Equilibria 191(1–2), 177–188.

An, C. L., Lim, W. J., Hong S. Y., Kim, E. J., Shin, E. C., Kim, M. K., Lee, J. R., Park, S. R., Woo, J. G., Lim, Y. P. & Yun, H. D. Analysis of bgl operon structure and characterization of β-Glucosidase from Pectobacterium carotovorum subsp. carotovorum LY34. Biosci. Biotechnol. Biochem. 68(11), 2270–2278.

Andreev, K. P. & Tsirlin, Y. A. 1954. Vapor-liquid equilibrium of the system furfural-methanol-water. Zhurnal Prikladnoi Khimii (Sankt-Petersburg, Russian Federation) 27, 402–412.

Andrić, P., Meyer, A. S., Jensen, P. A. & Dam-Johansen, K. 2010a. Effect and modeling of glucose inhibition and *In situ* glucose removal during enzymatic hydrolysis of pretreated wheat straw. Appl Biochem Biotechnol 160, 280–297.

Andrić, P., Meyer, A. S., Jensen, P. A. & Dam-Johansen, K. 2010b. Reactor design for minimizing product inhibition during enzymatic lignocellulose hydrolysis: I. Significance and mechanism of cellobiose and glucose inhibition on cellulolytic enzymes. Biotechnol. Adv. 28, 308–324.

Aristovich, V. Y., Lutugina, N. V., Malenko, Y. I. & Morachevskii, A. G. 1960. Liquid-vapor equilibriums in the process of rectification of the ternary system water-formic acid-acetic acid. Zhurnal Prikladnoi Khimii (Sankt-Petersburg, Russian Federation) 33, 2693–2698.

AspenTech. 2010. Process modelling tool for conceptual design, optimisation and performance monitoring of chemical processes.

Baez, A., Cho, K.-M. & Liao, J. C. 2011. High-flux isobutanol production using engineered Escherichia coli: a bioreactor study with in situ product removal. Appl Microbiol Biotechnol 90, 1681–1690.

Baker, J. O., Tatsumoto, K., Grohmann, K., Woodward, J., Wichert, J. M., Shoemaker, S. P. & Himmel, M. E. 1992. Thermal denaturation of Trichoderma reesei cellulases studied by differential scanning calorimetry and tryptophan fluorescence. Applied Biochemistry and Biotechnology 34/35, 217–231.

Bamberger, A., Sieder, G. & Maurer, G. 2004. High-pressure phase equilibrium of the ternary system carbon dioxide + water + acetic acid at temperatures from 313 to 353 K. Journal of Supercritical Fluids 32(1–3), 15–25.

Bansal, P., Hall, M., Realff, M. J., Lee, H. & Bommarius, A. S. 2009. Modeling cellulase kinetics on lignocellulosic substrates. Biotechnol. Adv. 27, 833–848.

Barker, J. A. & Henderson, D. 1967a. Perturbation theory and equation of state for fluids. II. A successful theory of liquids. The Journal of chemical physics 47, 4714.

Barker, J. A. & Henderson, D. 1967b. Perturbation theory and equation of state for fluids: The SquareWell potential. The Journal of chemical physics 47, 2856.

Baugh, K. D. & McCarty, P. L. 1988. Thermochemical pretreatment of lignocellulose to enhance methane fermentation: I. Monosaccharide and Furfurals hydrothermal decomposition and product formation rates. Biotechnology and Bioengineering 31, 50–61.

Baughman, A. C., Sharfstein, S. T. & Martin, L. L. 2011. A flexible state-space approach for the modeling of metabolic networks I: Development of mathematical methods, Metabolic Engineering 13(2), 125–137.

Bazaraa, M. S., Herali, H. D. & Shetty, C. M. 1993. Nonlinear programming: Theory and algorithms. John Wiley and Sons, New York, 2nd edition.

Becker, D., Braet, C., Brumer H. III., Clayssens, M., Divnes, C., Fagerström, B. R., Harris, M., Jones, T. A., Kleywegt, G. J., Koivula, A., Mahdi, S., Piens, K., Sinnott, M. L., Ståhlberg, J., Teeri, T. T., Underwood, M. & Wohlfahrt, G. 2001. Engineering of a glycosidase Family 7 cellobiohydrolase to more alkaline pH optimum: the pH behaviour of *Trichoderma reesei* Cel7A and its E223S/A224H/L225V/T226A/D262G mutant. Biochem. J. 356, 19–30.

Beckham, G., Bomble, Y., Bayer, E., Himmel, M. & Crowley, M. 2011. Applications of computational science for understanding enzymatic deconstruction of cellulose. Current Opinion in Biotechnology 22, 231–238.

Bezerra, R. M. F. & Dias, A. A. 2005. Enzymatic Kinetic of Cellulose Hydrolysis, Inhibition by Ethanol and Cellobiose. Applied Biochemistry and Biotechnology 126, 49–59.

Bhandari, N., Macdonald, D. G. & Bakhshi, N. N. 1984. Kinetic studies of corn stover saccharification using sulphuric acid. Biotechnology and Bioengineering 26, 320–327.

Blomberg, P. B. A. & Koukkari, P. S. 2009. The combination of transformed and constrained Gibbs energies. Math. Biosci. 220(2), 81–88.

Blomberg, P. B. A. & Koukkari, P. S. 2011. A systematic method to create reaction constraints for stoichiometric matrices. Comput. Chem. Eng. 35(7), 1238–1250.

Boerjan, W., Ralph, J. & Baucher, M. 2003. Lignin biosynthesis. Annual Review of Plant Biology 54(1), 519–546.

Bohlin, C., Olsen, S. N., Morant, M. D., Patkar, S., Borch, K. & Westh, P. A. 2010. Comparative study of activity and apparent inhibition of fungal β-Glucosidases. Biotechnol. Bioeng. 107, 943–952.

Börjesson, J., Engqvist, M., Sipos, B. & Tjerneld, F. 2007. Effect of poly(ethylene glycol) on enzymatic hydrolysis and adsorption of cellulase enzymes to pretreated lignocellulose. Enzyme and Microbial Technology 41, 186–195.

Branke, J., Deb, K., Miettinen, K. & Slowinski, R. (Eds.). 2008. Multiobjective optimization – Interactive and evolutionary approaches. Springer.

Bridgwater, A. V. & Cottam, M. 1992. Opportunities for biomass pyrolysis liquids production and upgrading. Energy & Fuels 6(2), 113–120.

Buhmann, M. D. 2003. Radial basis functions. Cambridge University Press.

Chang, C., Ma, X. & Cen, P. 2006. Kinetics of Levulinic Acid Formation from Glucose Decomposition at High Temperature. Chinese Journal of Chemical Engineering 14, 708–712.

Chang, C., Ma, X. & Cen, P. 2009. Kinetic Studies on Wheat Straw Hydrolysis to Levulinic Acid. Chinese Journal of Chemical Engineering 17, 835–839.

Chapman, W. G., Gubbins, K. E., Jackson, G. & Radosz, M. 1990. New reference equation of state for associating liquids. Industrial & Engineering Chemistry Research 29(8), 1709–1721.

Chen, H. & Jin, S. 2006. Effect of ethanol and yeast on cellulase activity and hydrolysis of crystalline cellulose. Enzyme and Microbial Technology 39, 1430–1432.

Collet, P. 2007. Genetic programming, in: Handbook of Research on Nature Inspired Computing for Economics and Management, Rennard, J.-P. (Ed.), Idea, Hershey, PA, 59–73.

COM. 2007. Commission decision of 21 December 2006 establishing harmonised efficiency reference values for separate production of electricity and heat in application of Directive 2004/8/EC of the European Parliament and of the Council. Official Journal of the European Union, L32, 183–188.

COM. 2010. Commission decision of 10 June 2010 on guidelines for the calculation of land carbon stocks for the purpose of Annex V to Directive 2009/28/EC. Official Journal of the European Union, L151, 19–41.

Conti, J. J., Othmer, D. F. & Gilmont, R. 1960. Composition of vapors from boiling binary solutions systems containing formic acid, acetic acid, water, and chloroform. Journal of Chemical and Data 5(3), 301–307.

Converse, A. O. & Optekar, J. D. 1993. A synergistic kinetic model for enzymatic cellulose hydrolysis compared to degree-of-synergism experimental results. Biotechnology and Bioengineering 42, 145–148.

Cortes, C. & Vapnik, V. 1995. Support vector networks. Machine Learning 20, 273–297.

COSMOlogic GmbH & Co. 2009. COSMOtherm, version C2.1, release 01.10. Leverkusen, Germany.

COSMO-SAC. COSMO-SAC-VT-2005 FORTRAN program and VT-2005 sigma profile database. 2010, from http://www.design.che.vt.edu/VT-Databases.html.

Cressie, N. 1990. The origins of kriging. Mathematical Geology 22, 197–202.

Cukor, P. M. & Prausnitz, J. M. 1972. Solubilities of gases in liquids at elevated temperatures. Henry's constants for hydrogen, methane, and ethane in hexadecane, bicyclohexyl, and diphenylmethane. The Journal of physical chemistry 76(4), 598–601.

Czernik, S. & Bridgwater, A. V. 2004. Overview of applications of biomass fast pyrolysis oil. Energy & Fuels 18(2), 590–598.

Dahl, S., Dunalewicz, A., Fredenslund, A. & Rasmussen, P. 1992. The MHV2 model: Prediction of phase equilibria at sub- and supercritical conditions. The Journal of Supercritical Fluids 5(1), 42–47.

DIPPR project 801 – full version 2005/2008/2009.

Djordjević, B. D., Šerbanović, S. P., Radović, I. R., Tasić, A. Ž. & Kijevčanin, M. L. J. 2007. Modelling of volumetric properties of binary and ternary mixtures by CEOS, CEOS/GE and empirical models. Journal of the Serbian Chemical Society 72(12), 1437–1463.

Domalski, E. S. & Hearing, E. D. 1993. Estimation of the thermodynamic properties of carbon-hydrogen-nitrogen-oxygen-sulfur-halogen compounds at 298.15 K. Journal of Physical and Chemical Reference Data 22(4), 805–1159.

Duarte, N., Herrgård, M. & Palsson, B. 2004. Reconstruction and validation of Saccharomyces cerevisiae iND750, a fully compartmentalized genome-scale metabolic model. Genome Res 14, 1298–1309.

EC. 2009. Directive 2009/28/EC on the European Parliament and of the Council of April 2009 on the promotion of the use of energy from renewable sources and amending and subsequently repealing Directives 2001/77/EC and 2003/30/EC. Official Journal of the European Union, L140, 16–62.

Ecoinvent. 2010. Data V2.2. Yeast paste from whey at fermentation.

Edwards, R., Larive, J., Mahieu, V. & Rouveirolles, P. 2007. Well-to-wheels analysis of future automotive fuels and powertrains in the European Context. WELL-to-WHEELS Report, Version 2c, March 2007. 140 p.

EN ISO 14040, 2006. Environmental management – Life cycle assessment – Principles and framework. International organization for standardization (ISO).

Eriksson, T., Karlsson, J. & Tjerneld, F. 2002. A Model Explaining Declining Rate in Hydrolysis of Lignocellulose Substrates with Cellobiohydrolase I (Cel7A) and Endoglucanase I (Cel7B) of *Trichoderma reesei*. Applied Biochemistry and Biotechnology 101, 41–60.

Esteghlalian, A., Hashimoto, A. G., Frenske, J. J. & Penner, M. H. 1997. Modelling and optimization of the dilute-sulfuric-acid pretreatment of corn stover, poplar and switchgrass. Bioresource Technology 59, 129–136.

Feng, W., van der Kooi, H. J. & de Swaan Arons, J. 2005. Application of the SAFT equation of state to biomass fast pyrolysis liquid. Chemical Engineering Science 60(3), 617–624.

Fitzpatrick, S. 1990. Lignocellulose Degradation to Furural and Levulinic Acid. U.S. Patent US4897497.

Fitzpatrick, S. 1997. Production of Levulinic Acid from Carbohydrate-Containing Materials. U.S. Patent US5608105.

Garrote, G., Domínguez, H. & Parajó, J. C. 2001. Kinetic modelling of corncob autohydrolysis. Process Biochemistry 36, 571–578.

Giri, B. K., Hakanen, J., Miettinen, K. & Chakraborti, N. 2011. Genetic Programming through Bi-objective Genetic Algorithms with Study of a Simulated Moving Bed Process involving Multiple Objectives. (*Submitted to Computers & Chemical Engineering.)*

Girisuta, B. 2007. Levulinic acid from lignocellulosic biomass. Dissertation, Rijksuniversiteit Groningen.

Gmehling, J., Li, J. & Fischer, K. 1997. Further development of the PSRK model for the prediction of gas solubilities and vapor-liquid -equilibria at low and high pressures II. Fluid Phase Equilibria 141(1–2), 113–127.

Goldberg, R. N. & Tewari, Y. B. 1989. Thermodynamic and Transport Properties of Carbohydrates and their Monophosphates: The Pentoses and Hexoses. J. Phys. Chem. Ref. Data 18, 809–880.

Gray, M. C., Converse, A. O. & Wyman, C. E. 2003. Sugar monomer and oligomer solubility. Applied Biochemistry and Biotechnology 105–108, 179–193.

Gregorich, E. G., Rochette, P., VandenBygaart, A. J. & Angers, D. A. 2005. Greenhouse gas contributions of agricultural soils and potential mitigation practices in Eastern Canada. Soil and Tillage Research 83(1), 53–72 with corrigendum in Soil and Tillage Research 94(1), 262–263.

Gross, J. & Sadowski, G. 2000. Application of perturbation theory to a hard-chain reference fluid: An equation of state for square-well chains. Fluid Phase Equilibria 168(2), 183–199.

Gross, J. & Sadowski, G. 2001. Perturbed-chain SAFT: An equation of state based on a perturbation theory for chain molecules. Industrial & Engineering Chemistry Research 40(4), 1244–1260.

Gross, J. & Sadowski, G. 2002a. Application of the perturbed-chain SAFT equation of state to associating systems. Industrial & Engineering Chemistry Research 41(22), 5510–5515.

Gross, J. & Sadowski, G. 2002b. Modeling polymer systems using the perturbed-chain statistical associating fluid theory equation of state. Industrial & Engineering Chemistry Research 41(5), 1084–1093.

Gross, J., Spuhl, O., Tumakaka, F. & Sadowski, G. 2003. Modeling copolymer systems using the perturbed-chain SAFT equation of state. Ind. Eng. Chem. Res. 42, 1266–1274.

Haar, L., Gallagher, J. S. & Kell, J. H. 1984. *NBS/NRC steam table*s. Washington: Hemisphere Publishing Corporation.

Hakanen, J., Kawajiri, Y., Miettinen, K. & Biegler, L. T. 2007. Interactive multi-objective optimization for simulated moving bed processes. Control & Cybernetics 36, 283–302.

Hammersley, J. 1960. Monte Carlo methods for solving multivariable problems. Proceedings of the New York Academy of Science 86, 844–874.

Harju, H. 2010. Aineensiirtoilmiöt guajakolin vetykäsittelyssä jalometallikatalyytillä. Aalto University. (In Finnish)

Hayden, J. G. & O'Connell, J. P. 1975. Generalized method for predicting second virial coefficients. Industrial & Engineering Chemistry Process Design and Development 14(3), 209–216.

Hayes, D. J. Fitzpatrick, S., Hayes, M. H. B. & Ross, J. R. 2006. The Biofine Process – Production of Levulinic Acid, Furfural and Formic Acid from Lignocellulosic Feedstocks. Biorefineries – Industrial Processes and Products. Wiley-VHC Verlag GmbH & Co., 139–164.

Haykin, S. 1998. Neural Networks: A Comprehensive Foundation, 2nd Ed. Prentice Hall.

Heiss-Blanquet, S., Zheng, D., Lopes Ferreira, N., Lapierre, C. &. Baumberger, S. 2011. Effect of pretreatment and enzymatic hydrolysis of wheat straw on cell wall composition, hydrophobicity and cellul ase adsorption. Bioresource Technology 102, 5938–5946.

Heric, E. L. & Rutledge, R. M. 1960a. Distribution of acetic and propionic acids between furfural and water. Journal of Chemical and Data 5(3), 272–274.

Heric, E. L. & Rutledge, R. M. 1960b. A method for the determination of tie-lines in ternary liquid systems. Canadian Journal of Chemical Engineering 38, 46–48.

Herrgard, M. J., Swainston, N., Dobson, P., Dunn, W. B., Arga, K. Y., Arvas, M., Bluthgen, N., Borger, S., Costenoble, R., Heinemann, M., Hucka, M., Le Novere, N., Li, P., Liebermeister, W., Mo, M. L., Oliveira, A. P., Petranovic, D., Pettifer, S., Simeonidis, E., Smallbone, K., Spasic, I., Weichart, D., Brent, R., Broomhead, D. S., Westerhoff, H. V., Kirdar, B., Penttilä, M., Klipp, E., Palsson, B. O., Sauer, U., Oliver, S. G., Mendes, P., Nielsen, J. & Kell, D. B. 2008. A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology. Nat. Biotechnol. 26, 1155–1160.

Hjersted. J. L & Henson, M. A. Optimization of fed-batch Saccharomyces cerevisiae fermentation using dynamic flux balance models. Biotechnol Prog. 2006 Sep-Oct;22(5):1239-48.

Hodge, D., Karim, M., Schell, D. & McMillan, J. 2009. Model-based fed-batch for high-solids enzymatic cellulose hydrolysis. Appl Biochem Biotechnol 152, 88–107.

Holderbaum, T. & Gmehling, J. 1991. PSRK: A group contribution equation of state based on UNIFAC. Fluid Phase Equilibria 70(2–3), 251–265.

Hollenberg, J.-L. & Hall, D. O. 1983. Hydration numbers by near-infrared spectrophotometry. 2. Sugars. J. Phys. Chem. 87(4), 695–696.

Holzhütter, H.-G. 2004. The principle of flux minimization and its application to estimate stationary fluxes in metabolic networks. Eur. J. Biochem. 271, 2905–2922.

Horiuti, I. & Polanyi, M. 1934. Exchange reactions of hydrogen on metallic catalysts. Transactions of the Faraday Society 30, 1164–1172.

Horstmann, S., Fischer, K. & Gmehling, J. 2000. PSRK group contribution equation of state: Revision and extension III. Fluid Phase Equilibria 167(2), 173–186.

Horstmann, S., Jabłoniec, A., Krafczyk, J., Fischer, K. & Gmehling, J. 2005. PSRK group contribution equation of state: Comprehensive revision and extension IV, including critical constants and α-function parameters for 1000 components. Fluid Phase Equilibria 227(2), 157–164.

Huang, L., Forsberg, C. W. & Thomas, D. Y. 1988. Purification and characterization of a chloride-stimulated cellobiosidase from Bacteroides succinogenes S85. Journal of Bacteriology 170(7), 2923–2932.

Hypercube, I. 2002. Molecular modeling system, hyperchemHypercube, Inc.

Igarashi, K., Uchihashi, T., Koivula, A., Wada, M., Kimura, S., Okamoto, T., Penttilä, M., Ando, T. & Samejima, S. 2011. Traffic jams reduce hydrolytic efficiency of cellulase on cellulose surface. Science 333, 1279–1282.

Igarashi, K., Wada, M., Hori, R. & Samejima, M. 2006. Surface density of cellobiohydrolase on crystalline celluloses. A critical parameter to evaluate enzymatic kinetics at a solid–liquid interface. FEBS Journal 273, 2869–2878.

Jacobsen, S. E. & Wyman, C. E. 2000. Cellulose and hemicellulose hydrolysis models for application to curre nt and novel pretreatment processes. Applied Biochemistry and Biotechnology 84–86, 81–96.

Jäger, G., Wu, Z., Garschhammer, K., Engel, P., Klement, T., Rinaldi, R., Spiess, A. & Büchs, J. 2010. Practical screening of purified cellobiohydrolases and endoglucanases with a-cellulose and specification of hydrodynamics. Biotechnology for Biofuels 3(18), 12 p.

Jankowski, M.D., Henry, C.S., Broadbelt, L.J. & Hatzimanikatis, V. 2008. Group contribution method for thermodynamic analysis of complex metabolic networks. Biophys J. 95(3), 1487–99.

Jansson, P.-E. & Karlberg, L. 2001. Coupled heat and mass transfer model for soil-plant-atmosphere systems. Royal Inst. Technol., Dept. Civil and Environ. Eng., Stockholm. ftp://www.lwr.kth.se/CoupModel/CoupModel.pdf.

Jeoh, T., Baker, J. O., Ali, M. K., Himmel, M. E. & Adney, W. S. 2005. β-D-Glucosidase reaction kinetics from isothermal titration microcalorimetry, Analytical Biochemistry 347, 244–253.

Joback, K. G. & Reid, R. C. 1987. Estimation of pure-component properties from group contributions. Chemical Engineering Communications 57(1–6), 233–243.

Jog, P. K. & Chapman, W. G. 1999. Application of wertheim's thermodynamic perturbation theory to dipolar hard sphere chains. Molecular Physics 97(3), 307–319.

Jørgensen, H., Vibe-Pedersen, J., Larsen, J. & Felby, C. 2007. Liquefaction of lignocellulose at high-solids concentrations. Biotechnol. Bioeng. 96, 862–870.

Kadam, K., Rydholm, E. & McMillan, J. 2004. Development and validation of a kinetic model for enzymatic saccharification of lignocellulosic biomass. Biotechnol. Prog. 20, 698–705.

Karboune, S., Geraert, P.-A. & Kermasha, S. 2008. Characterization of selected cellulolytic activities of multi-enzymatic complex system from Penicillium funiculosum. J. Agric. Food Chem. 56, 903–909.

Kawajiri, Y. & Biegler, L. T. 2006. Optimization strategies for simulated moving bed and powerfeed processes. AIChE Journal 52, 1343–1350.

Keskinen, K. I. & Aittamaa, J. 2010. User Manual of Program Flowbat.

Kim, D. W. & Hong, Y. G. 2001. Description of Cellobiohydrolases Ce16A and Ce17A from Trichoderma reesei Using Langmuir-type Models. Biotechnol. Bioprocess Eng. 6, 89–94.

Klamt, A. 1995. Conductor-like screening model for real solvents: A new approach to the quantitative calculation of solvation phenomena. Journal of Physical Chemistry 99(7), 2224–2235.

Klamt, A. 2005. COSMO-RS: From quantum chemistry to fluid phase thermodynamics and drug design.

Klamt, A., Krooshof, G. J. P. & Taylor, R. 2002. COSMOSPACE: Alternative to conventional activity-coefficient models. AICHE Journal 48(10), 2332–2349.

Knovel critical tables. 2008. http://www.knovel.com/web/portal/browse/display?_EXT_KNOVEL_DISPLAY_bookid=761&VerticalID=0.

Koivula, A., Ruohonen, L., Wohlfahrt, G., Reinikainen, T., Teeri, T. T., Piens, K., Claeyssens, M., Weber, M., Vasella, A., Becker, D., Sinnott, M. L., Zou, J.-Y., Kleywegt, G. J., Szardenings, M., Ståhlberg, J. & Jones, T. A. 2002.The Active Site of Cellobiohydrolase Cel6A from Trichoderma reesei: The Roles of Aspartic Acids D221 and D175. J. Am. Chem. Soc. 124, 10015–10024.

Kontogeorgis, G. M., Voutsas, E. C., Yakoumis, I. V. & Tassios, D. P. 1996. An equation of state for associating fluids. Industrial and Engineering Chemistry Research 35(11), 4310–4318.

Koukkari, P. (Ed.). 2009. Advanced Gibbs Energy Methods for Functional Materials and Processes – ChemSheet 1999–2009. VTT, Espoo. VTT Research Notes 2506. 145 p. http://www.vtt.fi/inf/pdf/tiedotteet/2009/T2506.pdf.

Koukkari, P., Pajarre, R. & Blomberg, P. 2011. Reaction rates as virtual constraints in Gibbs energy minimization. Pure Appl. Chem. 83(5), 1063–1074.

Koukkari, P., Pajarre, R. & Pakarinen, H. 2002. Modeling of the ion exchange in pulp suspensions by Gibbs energy minimization. J. Solution Chem. 31(8), 627–638.

Kristensen, J. B., Börjesson, J., Bruun, M. H., Tjerneld, F. & Jørgensen, H. 2007. Use of surface active additives in enzymatic hydrolysis of wheat straw lignocellulose. Enzyme and Microbial Technology 40, 888–895.

Kristensen, J. B., Felby, C. & Jørgensen, H. 2009. Yield-determining factors in high-solids enzymatic hydrolysis of lignocellulose. Biotechnology for Biofuels 2(11), 10 p.

Kushner, T. M., Tatsievskaya, G. I., Irzun, V. A., Volkova, L. V. & Serafimov, L. A. 1966. Liquid-vapor phase equilibrium in the water-formic acid-acetic acid system under 50 and 200 mm. hg pressure. Zhurnal Fizicheskoi Khimii 40(12), 3010–3017.

Kwon, I., Ekino, K., Goto, M. & Furukawa, K. 1999. Heterologous expression and characterization of endoglucanase I (EGI) from Trichoderma viride HK-75. Biosci. Biotechnol. Biochem. 63(10), 1714–1720.

Langford, R. E. & Heric, E. L. 1972. Furfural-water-formic acid system at 25° and 35°C. Journal of Chemical and Engineering Data 17(1), 87–89.

Lee, S., Phalakornkule, C., Domach, M. M. & Grossmann, I. E. 2000. Comput. Chem. Eng. 24, 711–716.

Leppävuori, J., Diaz Domach, M. S. & Biegler, L. 2011. Parameter estimation in batch bioreactor simulation using metabolic models: Sequential solution with direct sensitivities. Ind. Eng. Chem. Res. 50, 12080–1209.

Li, S., Wen, J. & Jia, X. 2011. Engineering Bacillus subtilis for isobutanol production by heterologous Ehrlich pathway construction and the biosynthetic 2-ketoisovalerate precursor pathway overexpression. Appl Microbiol Biotechnol 91, 577–589.

Lin, S. & Sandler, S. I. 2002. A priori phase equilibrium prediction from a segment contribution solvation model. Industrial & Engineering Chemistry Research 41(5), 899–913.

Linder, M., Nevanen, T. & Teeri, T.T. 1999. Design of a pH-dependent cellulose-binding domain. FEBS Letters 447, 13–16.

Linder, M. & Teeri, T. 1996. The cellulose-binding domain of the major cellobiohydrolase of Trichoderma reesei exhibits true reversibility and a high exchange rate on crystalline cellulose. Proc. Natl. Acad. Sci. 93, 122251–12255.

Lindfors, C. 2008. Fractionation of bio oil. (Master, Helsinki University of Technology, Industrial Chemistry Laboratory).

Luong, J. H. T. 1985. Kinetics of ethanol inhibition in alcohol fermentation. Biotechnology and Bioengineering 27, 280–285.

Lylykangas, M. 2004. Kinetic modeling of liquid-phase hydrogenation reactions. Helsinki University of Technology.

Ma, A., Hu, Q., Qu, Y., Bai, Z., Liu, W. & Zhuang, G. 2008. The enzymatic hydrolysis rate of cellulose decreases with irreversible adsorption of cellobiohydrolase I. Enzyme and Microbial Technology 42, 543–547.

Mahadevan, R. & Schilling, C. H. 2003. The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. Metab. Eng. 5, 264–276.

Mäkinen, T., Soimakallio, S., Paappanen, T., Pahkala, K. & Mikkola , H. 2006. Greenhouse gas balances and new business opportunities for biomass-based transportation fuels and agrobiomass in Finland. VTT, Espoo. VTT Research Notes 2357. 134 p. + app. 19 p. http://www.vtt.fi/inf/pdf/tiedotteet/2006/T2357.pdf. (In Finnish.)

Malhi, S. S., Lemke, R., Wang, Z. H., Chabra, S. & Baldev, S. 2006. Tillage, nitrogen and crop residue effects on crop yield, nutrient uptake, soil quality, and greenhouse gas emissions. Soil and Tillage Research 90(1–2), 171–183.

Malhi, S. S. and Lemke, R. 2007. Tillage, crop residue and N fertilizer effects on crop yield, nutrient uptake, soil quality and nitrous oxide gas emissions in a second 4-yr rotation cycle, Soil and Tillage Research, Volume 96, Issue 1–2, October 2007, Pages 269–283.

Marrero, J. & Gani, R. 2001. Group-contribution based estimation of pure component properties. Fluid Phase Equilibria 183–184, 183–208.

Marrero-Morejon, J. & Pardillo-Fontdevila, E. 1999. Estimation of pure compound properties using group-interaction contributions. AICHE Journal 45(3), 615–621.

Maurelli, L., Giovane, A., Esposito, A., Moracci, M., Fiume, I., Rossi, M. & Morana, A. 2008. Evidence that the xylanase activity from Sulfolobus solfataricus Oα

is encoded by the endoglucanase precursor gene (sso1354) and characterization of the associated cellulase activity. Extremophiles 12, 689–700.

McKibbins, S. W., Harris, J. F. & Saeman, J. F. 1962. Kinetics of the Acid Catalyzed Conversion of Glucose to 5-Hydroxymethyl-2-Furaldehyde and Levulinic Cid. Forest Products Journal 12, 17–23.

Michelsen, M. L. 1990. A modified huron-vidal mixing rule for cubic equations of state. Fluid Phase Equilibria 60(1–2), 213–219.

Mullins, E., Oldland, R., Liu, Y. A., Wang, S., Sandler, S. I. & Chen, C. 2006. Sigma-profile database for using COSMO-based thermodynamic methods. Industrial and Engineering Chemistry Research 45(12), 4389–4415.

Murabito, E., Simeonidis, E., Smallbone, I. &. Swinton, J. 2009. Vertical systems biology: from DNA to flux and back. Theor. Biol. 260, 445–452.

Murayama, Y. 1961. Vapor-liquid equilibriums of acetic acid-formic acid-water. Nippon Kagaku Zasshi 82, 550–554.

Nannoolal, Y., Rarey, J. & Ramjugernath, D. 2007. Estimation of pure component properties: Part 2. Estimation of critical property data by group contribution. Fluid Phase Equilibria 252(1–2), 1–27.

Nannoolal, Y., Rarey, J. & Ramjugernath, D. 2008. Estimation of pure component properties: Part 3. Estimation of the vapor pressure of non-electrolyte organic compounds via group contributions and group interactions. Fluid Phase Equilibria 269(1–2), 117–133.

Nannoolal, Y., Rarey, J., Ramjugernath, D. & Cordes, W. 2004. Estimation of pure component properties: Part 1. Estimation of the normal boiling point of non-electrolyte organic compounds via group contributions and group interactions. Fluid Phase Equilibria 226, 45–63.

Nidetzky, B., Steiner, W. & Cleyssens, M. 1994. Cellulose hydrolysis by the cellulases from Trichoderma reesei: adsorptions of two cellobiohydrolases, two endoglucanases and their core proteins on filter paper and their relation to hydrolysis. Biochem. J. 303, 817–823.

Nielsen, P. H., Oxenbøll, K. M. & Wenzel, H. 2007. Cradle-to-Gate Environmental Assessment of Enzyme Products Produced Industrially in Denmark by Novozymes A/S. International Journal of Life Cycle Assessment 12(6), 432–438.

Nocedal, J. & Wright, S. J. 2000. Numerical Optimization. Springer.

Nothnagel, K. H., Abrams, D. S. & Prausnitz, J. M. 1973. Generalized correlation for fugacity coefficients in mixtures at moderate pressures. Application of chemical theory of vapor imperfections. Industrial & Engineering Chemistry Process Design and Development 12(1), 25–35.

Oasmaa, A., Kuoppala, E. & Solantausta, Y. 2003. Fast pyrolysis of forestry residue. 2. physicochemical composition of product liquid. Energy & Fuels 17(2), 433–443.

Oja, V. & Suuberg, E. M. 1999. Vapor pressures and enthalpies of sublimation of D-glucose, D-xylose, cellobiose, and levoglucosan. Journal of Chemical and Engineering Data 44(1), 26–29.

Palmqvist, E. & Hahn-Hägerdal, B. 2000. Fermentation of lignocellulosic hydrolysates. II: inhibitors and mechanisms of inhibition. Bioresource Technology 7, 25–33.

Palonen, H., Tjerneld, F., Zacchi, G. & Tenkanen, M. 2004. Adsorption of *Trichoderma reesei* CBH I and EG II and their catalytic domains on steam pretreated softwood and isolated lignin. Journal of Biotechnology 107, 65–72.

Park, J. M., Kim, T. Y. & Lee, S. Y. 2009. Constraints-based genome-scale metabolic simulation for systems metabolic engineering. Biotechnology Advances 27(6), 979–988.

Pegoraro, M. & Guglielmi, G. 1955. Ternary system furfural-water-acetic acid. solubility curves and the corresponding liquid-phase equilibriums at various temperatures. Chimica e l'Industria (Milan, Italy) 37, 1035–1038.

Peneloux, A., Rauzy, E. & Freze, R. 1982. A consistent correction for redlich-kwong-soave volumes. Fluid Phase Equilibria 8, 7–23.

Pereira, F. B., Guimarães, P. M. R., Teixeira, J. A. & Domingues, L. 2011. Robust industrial Saccharomyces cerevisiae strains for very high gravity bio-ethanol fermentations. Journal of Bioscience and Bioengineering 112(2), 130–136.

Peres, A. M. & Macedo, E. A. 1997. Phase equilibria of -glucose and sucrose in mixed solvent mixtures: Comparison of UNIQUAC 1-based models. Carbohydrate research 303(2), 135–151.

Pitkänen, J.-P. 2005. Impact of xylose and mannose on central metabolism of yeast *Saccharomyces cerevisiae*. Espoo 2005, Helsinki University of Technology. ISBN 951-22-7893-6 (printed), ISBN 951-22-7894-4 (pdf), online: http://lib.tkk.fi/Diss/2005/isbn9512278944.

Poling, B. E., Prausnitz, J. M. & O'Connell, J. P. 2001. In: Poling, B. E., Prausnitz, J. M. & O'Connell, J. P. (Eds.). Properties of gases and liquids (5th edition) McGraw-Hill.

Prausnitz, J. M., Lichtenhaler, R. N. & Azevedo, E. G. D. 1986. *Molecular thermodynamics of fluid-phase equilibria* (2. ed.). Englewood Cliffs, NJ: Prentice-Hall.

Reese, E. T. & Ryu, D. Y. 1980. Shear inactivation of cellulase of Trichoderma reesei. Enzyme Microb. Technol. 2, 239–240.

Ruttersmith, L. D. & Daniel, R. M. 1991. Thermostable cellobiohydrolase from the thermophilic eubacterium Thermotoga sp. strain FjSS3-B.1. Biochem. J. 277, 887–890.

Saeman, J. F. 1945. Kinetics of wood saccharification – Hydrolysis of cellulose and decomposition of sugars in dilute acid at high temperature. Ind. Eng. Chem. Res. 37, 43–52.

Sako, T., Sugeta, T., Nakazawa, N., Otake, K., Sato, M. & Ishihara, K. 1995. High pressure vapor-liquid and vapor-liquid-liquid equilibria for systems containing supercritical carbon dioxide, water and furfural. Fluid Phase Equilibria 108(1–2), 293–303.

Salo, T. 2007. E-mails in February 2007 (Salo – Soimakallio) and August 2007 (Salo – Tsupari).

Sandler, S. I. 2006. Chemical, biochemical, and engineering thermodynamics (4th ed.). Hoboken, N.J: John Wiley.

Sanghvi, R. & Yalkowsky, S. H. 2006. Estimation of the normal boiling point of organic compounds. Industrial and Engineering Chemistry Research 45(8), 2856–2861.

Sawistowski, H. & Pilavakis, P. A. 1982. Vapor-liquid equilibrium with association in both phases. multicomponent systems containing acetic acid. Journal of Chemical and Engineering Data 27(1), 64–71.

Schafer, A., Klamt, A., Sattel, D., Lohrenz, J. C. W. & Eckert, F. 2000. COSMO implementation in TURBOMOLE: Extension of an efficient quantum chemical code towards liquid systems. Physical Chemistry Chemical Physics 2(10), 2187–2193.

Shlomi, T., Benyamini, T., Gottlieb, E., Sharan, R. & Ruppin, E. 2011. Genome-scale metabolic modeling elucidates the role of proliferative adaptation in causing the Warburg effect. PLoS Comput. Biol. 7(3), e1002018.

Sipos, B., Benkő, Z., Dienes, D., Réczey, K., Viikari, L. & Siika-aho, M. 2010. Characterisation of Specific Activities and Hydrolytic Properties of Cell-Wall-Degrading Enzymes Produced by Trichoderma reesei Rut C30 on Different Carbon Sources, Appl Biochem Biotechnol 161, 347–364.

Sitaraman, R. 1963. Generalized equation for diffusion in liquids. Journal of Chemical and Engineering Data 8(2), 198–201.

Smith, K. M., Cho, K.-M. & Liao, J. C. 2010. Engineering Corynebacterium glutamicum for isobutanol production. Appl Microbiol Biotechnol 87, 1045–1055.

Soave, G. 1972. Equilibrium constants from a modified redlich-kwong equation of state. Chemical Engineering Science 27(6), 1197–1203.

Soimakallio, S. & Saikku, L. 2012. CO_2 emissions attributed to annual average electricity consumption in OECD countries. Manuscript submitted.

Song, H.-S., Morgan, J. A. & Ramkrishna, D. 2009. Systematic development of hybrid cybernetic models: Application to recombinant yeast co-consuming glucose and xylose. Biotechnol. Bioeng. 103, 984–1002.

Statistics Finland. 2011. Fuel classification and emission coefficients 2011. Available: http://tilastokeskus.fi/tup/khkinv/khkaasut_polttoaineluokitus.html.

Stein, S. E. & Brown, R. L. 1994. Estimation of normal boiling points from group contributions. Journal of chemical information and computer sciences 34(3), 581–587.

Suchy, M., Linder, M., Tammelin, T., Campbell, J. M., Vuorinen, T. & Kontturi, E. 2011. Quantitative Assessment of the Enzymatic Degradation of Amorphous Cellulose by Using a Quartz Crystal Microbalance with Dissipation Monitoring, Langmuir 27, 8819–8828.

Suurnäkki, A., Tenkanen, M., Siika-aho, M., Niku-Paavola, M.-L., Viikari, L. & Buchert, J. 2000. Trichoderma reesei cellulases and their core domains in the hydrolysis and modification of chemical pulp, Cellulose 7, 189–209.

Tewari, Y. B., Lang, B. E., Decke, S. R. & Goldberg, R. N. 2008. Thermodynamics of the hydrolysis reactions of 1,4-β-D-xylobiose, 1,4- β -D-xylotriose, D-cellobiose, and D-maltose. J. Chem. Thermodynamics 40, 1517–1526,

Tohmola, N., Ahtinen, J., Pitkänen, J.-P., Parviainen, V., Joenväärä, S., Hautamäki, M., Lindroos, P., Mäkinen, J. & Renkonen, R. 2011. On-line high performance liquid chromatography measurements of extracellular metabolites in an aerobic batch yeast (Saccharomyces cerevisiae) culture. Biotech Bioproc Eng. 16, 264–272.

Törn, A. & Zilinskas, A. 1989. Global Optimization. Springer-Verlag.

Tsirlin, Y. A. 1969. Liquid-vapor phase equilibrium in a water-formic acid-furfural three-component system. Sb. Tr. Vses. Nauch.-Issled. Inst. Gidroliza Rast. Mater. 18, 190–203.

Tsupari, E., Monni, S., Tormonen, K., Pellikka, T. & Syri, S. 2007. Estimation of annual CH_4 and N_2O emissions from fluidised bed combustion: An advanced measurement-based method and its application to Finland. International Journal of Greenhouse Gas Control 1, 289–297.

Tyn, M. T. 1975. Diffusion coefficients in dilute binary liquid mixtures. [Electronic version]. Journal of Chemical and Engineering Data 20(1), 106–109.

van Zyl, J. M., van Rensburg, E., van Zyl, W. H., Harms, T. M. & Lynd, L. R. 2011. A Kinetic model for simultaneous saccharification and fermentation of avicel with *Saccharomyces cerevisiae.* Biotechnol. Bioeng. 108, 924–933.

Vane, L. M. 2008. Separation technologies for the recovery and dehydration of alcohols from fermentationbroths. Biofuels, Bioprod. Bioref. 2, 553–588.

Vapnik, V. N. 1998. Statistical learning theory. John Wiley & Sons, New York.

von Weymarn, N. (Ed.). 2007. Bioetanolia maatalouden selluloosavirroista. Espoo, VTT. VTT Tiedotteita – Research Notes 2412. http://www.vtt.fi/inf/pdf/tiedotteet/2007/T2412.pdf. (In Finnish)

Voutsas, E., Magoulas, K. & Tassios, D. 2004. Universal mixing rule for cubic equations of state applicable to symmetric and asymmetric systems: Results with the peng-robinson equation of state. Industrial and Engineering Chemistry Research 43(19), 6238–6246.

Wang, W., Kang, L., Wei, H., Arora, R. & Lee, Y. Y. 2011. Study on the decreased sugar yield in enzymatic hydrolysis of cellulosic substrate at high solid loading. Appl Biochem Biotechnol 164, 1139–1149.

Weise, T. 2009. Global Optimization book: http://www.itweise.de/projects/book.pdf.

Wertheim, M. S. 1983a. Fluids with highly directional attractive forces. I. statistical thermodynamics. Journal of Statistical Physics 35(1–2), 19–34.

Wertheim, M. S. 1983b. Fluids with highly directional attractive forces. II. thermodynamic perturbation theory and integral equations. Journal of Statistical Physics 35(1–2), 35–47.

Wertheim, M. S. 1985a. Fluids with highly directional attractive forces. III. multiple attraction sites. Journal of Statistical Physics 42(3–4), 459–476.

Wertheim, M. S. 1985b. Fluids with highly directional attractive forces. IV. equilibrium polymerization. Journal of Statistical Physics 42(3–4), 477–492.

Westerhof, R. J. M., Kuipers, N. J. M., Kersten, S. R. A. & van Swaaij, Wim, P. M. 2007. Controlling the water content of biomass fast pyrolysis oil. Industrial & Engineering Chemistry Research 46(26), 9238–9247.

Wilke, C. R. & Chang, P. 1955. Correlation of diffusion coefficients in dilute solutions. AIChE Journal 1(2), 264–270.

Wisniak, J. & Tamir, A. 1977. Vapor-liquid equilibria in the ternary systems water-formic acid-acetic acid and water-acetic acid-propionic acid. Journal of Chemical and Engineering Data 22(3), 253–260.

Wohlfahrt, G., Pellikka, T., Boer, H., Teeri, T. T. & Koivula, A. 2003. Probing pH-dependent functional elements in proteins: Modification of carboxylic acid pairs in *Trichoderma reesei* cellobiohydrolase Cel6A. Biochemistry 42, 10095–10103.

Wolbach, J. P. & Sandler, S. I. 1998. Using molecular orbital calculations to describe the phase behavior of cross-associating mixtures. Industrial & Engineering Chemistry Research 37(8), 2917–2928.

Yaws. 2010. Thermophysical properties of chemicals and hydrocarbons (electronic edition). Knovel.from http://www.knovel.com/web/portal/browse/display?_EXT_KNOVEL_DISPLAY_bookid=2906.

Yoon, J., Chun, M., Hong, W. & Lee, H. 1993. High-pressure phase equilibria for carbon dioxide-methanol-water system: Experimental data and critical evaluation of mixing rules. Industrial and Engineering Chemistry Research 32(11), 2881–2887.

Zhang, Y., Xu, J.-L., Xu, H.-J., Yuan, Z.-H. & Guo, Y. 2010. Cellulase deactivation based kinetic modeling of enzymatic hydrolysis of steam-exploded wheat straw. Bioresource Technology 101, 8261–8266.

Zhang, Y.-H. P. & Lynd, L.R. 2004. Toward an Aggregated Understanding of Enzymatic Hydrolysis of Cellulose: Noncomplexed Cellulase Systems. Biotechnology and Bioengineering 88(7), 797–824.

Zhang, Y.-H.P., Ding, S.-Y., Mielenz, J. R., Cui, J.-B., Elander, R. T., Laser, M., Himmel, M., McMillan, J.R. & Lynd, L.R. 2007. fractionating recalcitrant lignocellulose at modest reaction conditions. Biotechnol. Bioeng. 97, 214–223.

Zheng, Y., Pan, Z., Zhang, R. & Jenkins, B. 2009. Kinetic modeling for enzymatic hydrolysis of pretreated creeping wild ryegrass. Biotechnol. Bioeng. 102, 1558–1569.

Zhu, Z., Sathitsuksanoh, N., Vinzant, T., Schell, D. J., McMillan, J. D. & Zhang, Y.-H. P. 2009. Comparative study of corn stover pretreated by dilute acid and cellulose solvent-based lignocellulose fractionation: Enzymatic hydrolysis, supramolecular structure, and substrate accessibility. Biotechnol. Bioeng. 103, 715–724.

Appendix A: Binary interaction parameters for PC-SAFT, SRK model, parameter fitting accuracy, experimental data used for the fitting

1st comp	2nd comp	PC-SAFT k_ij	SRK k_ij	PC-SAFT error *	SRK error*	PSRK error *	# points	Quality	T ange (K)		p range (kPa)		Source
1,2-Ethanediol	Acetic acid	-0.1343	-0.0648	39.8	46.0	122.7	15	xTp	363.2	363.2	1.4	39.7	Schmid B. *et al.,* Fluid Phase Equilib., 2007, 258, 115–124
1,2-Ethanediol	m-Cresol	-0.0350	-0.0080	18.2	10.6	5.1	13	xyTp	468.2	475.6	101.3	101.3	Othmer D. F. *et al.,* Ind. Chem. Chem., 1949, 41, 572
1,2-Ethanediol	Methane	0.0431	0.1653	135.6	123.8	204.1	6	xTp	398.2	398.2	135.0	18170.0	Jou F. -Y. *et al.,* The Canadian J Chem Eng., 1994, 72, 130–133
1,2-Ethanediol	p-Cresol	-0.0343	0.0096	18.5	8.0	4.5	13	xyTp	467.9	475.2	101.3	101.3	Othmer D. F. *et al.,* Ind. Chem. Chem., 1949, 41, 572
Benzaldehyde	m-Cresol	-0.0195	-0.0432	7.0	6.9	20.0	3	xyTp	389.2	409.9	13.3	13.3	Vosrikova V. N. *et al.,* Khim. Prom. (Moscow), 1978, 5, 345
Benzaldehyde	o-Cresol	-0.0135	-0.0300	24.0	25.0	25.8	7	xyTp	409.5	415.9	26.7	26.7	Vostrikova V. N. *et al.,* Zh. Prikl. Khim. (Leningrad), 1980, 53, 2763
Benzaldehyde	p-Cresol	-0.0145	-0.0300	7.3	12.7	10.5	4	xyTp	390.3	408.1	13.3	13.3	Vostrikova V. N. *et al.,* Zh. Prikl. Khim., 1980, 53, 2118
Benzaldehyde	Phenol	-0.0185	-0.0488	12.6	19.8	13.2	6	xyTp	388.3	395.4	13.3	13.3	Vostrikova V. N. *et al.,* Zh. Prikl. Khim., 1978, 51, 211
Benzene	Acetic acid	-0.0574	0.0810	43.3	26.0	20.0	17	xyTp	352.8	391.4	101.3	101.3	Haughton C. O., Brit. Chem. Eng., 1967, 12, 1102
Benzene	Benzaldehyde	0.0079	0.0145	25.4	24.5	28.4	21	xTp	413.7	413.7	35.5	480.9	Nienhaus B. *et al.,* J. Chem. Eng. Data, 1998, 43, 941–948
Benzene	Cyclohexane	0.0145	0.0216	2.1	1.5	4.2	11	xyTp	423.2	423.2	58.2	604.3	Wisniewska B. *et al.,* Fluid Phase Equilib., 1993, 86, 173
Benzene	Cyclohexene	0.0094	0.0167	1.5	2.5	5.7	15	xyTp	352.1	355.3	98.7	98.7	Harrison J. M. *et al.,* Ind Eng Chem., 1946, 38, 117
Benzene	m-Cresol	0.0059	0.0091	37.3	40.0	50.1	11	xTp	353.3	474.7	101.3	101.3	Piatti L., Z. Phys. Chem., 1930, 152, 36
Benzene	Methylcyclo-	0.0120	0.0118	13.1	13.2	14.8	7	xTp	353.2	353.2	53.6	101.0	Van Pham S., Thesis Leipzig, 1983

1st comp	2nd comp	PC-SAFT k_ij	SRK k_ij	PC-SAFT error *	SRK error*	PSRK error *	# points	Quality	T ange (K)		p range (kPa)		Source
Benzene	Phenol	-1.0683	-2.1148	1094.9	1068.5	1834.2	17	xTp	413.6	413.6	30.2	480.3	Nienhaus B. *et al.,* J. Chem. Eng. Data, 1998, 43, 941–948
Benzene	Toluene	0.0024	0.0048	6.3	5.0	7.0	11	xTp	393.2	393.2	129.7	299.0	Schmidt G. C., Z. Phys. Chem., 1926, 121, 221
Benzene	Water	-0.0945	-0.1941	1588.4	1580.2	1966.3	4	xTp	298.0	473.0	5000.0	5000.0	Mathis J. *et al.,* J. Chem. Eng. Data, 2004, 49, 1269–1272
Carbon dioxide	1,2-Ethanediol	-0.0525	-0.1715	1120.7	991.0	1114.8	9	xTp	398.2	398.2	960.0	38400.0	Zheng D.-Q. *et al.,* Fluid Phase Equilib., 1999, 155, 277–286
Carbon dioxide	Acetic acid	0.0335	0.0195	95.3	104.8	104.7	7	xyTp	353.2	353.2	2100.0	11100.0	Bamberger A. *et al.,* Journal of Supercritical Fluids, 2000, 17, 97–110
Carbon dioxide	Benzaldehyde	0.0836	0.0466	67.7	36.6	100.6	9	xyTp	393.2	393.2	6170.0	22340.0	Walther D., Maurer G., Ber. Bunsen-Ges. Phys. Chem., 1992, 96, 981–988
Carbon dioxide	Benzene	0.1172	0.1135	79.0	68.7	81.9	16	xyTp	344.3	344.3	6895.0	10920.0	Nagarajan N., Robinson R. L. Jr., J. Chem. Eng. Data, 1987, 32, 369–371
Carbon dioxide	Catechol	0.0391	-0.0419	30.3	23.2	460.2	5	xyTp	473.2	473.2	1013.3	5066.3	Yau J., Tsal F., J. Chem. Eng. Data, 1992, 37, 141–143
Carbon dioxide	Cyclohexane	0.1488	0.1288	59.7	39.2	123.7	14	xyTp	344.3	344.3	6870.0	10930.0	Nagarajan N., Robinson R. L. Jr., J. Chem. Eng. Data, 1987, 32, 369–371
Carbon dioxide	Ethanol	0.0455	0.0977	162.5	64.7	63.3	11	xyTp	344.8	344.8	800.0	11930.0	Joung S. N. *et al.,* Fluid Phase Equilibria, 2001, 185, 219–230
Carbon dioxide	Ethyl acetate	-0.0339	-0.0617	121.1	141.5	247.3	7	xyTp	393.0	393.0	5060.0	11970.0	Tian Y.-L. *et al.,* J. Chem. Eng. Data, 2004, 49, 1554–1559
Carbon dioxide	Guaiacol	0.0834	0.0512	66.0	57.5	344.5	10	xyTp	393.2	393.2	2000.0	20000.0	Lee M. *et al.,* Fluid Phase Equilibria, 1999, 162, 211–224
Carbon dioxide	Methanol	0.0167	0.0685	95.5	55.7	72.5	13	xyTp	342.8	342.8	670.0	12390.0	Joung S. N. *et al.,* Fluid Phase Equilibria, 2001, 185, 219–230

1st comp	2nd comp	PC-SAFT k_ij	SRK k_ij	PC-SAFT error *	SRK error*	PSRK error *	# points	Quality	T ange (K)		p range (kPa)		Source
Carbon dioxide	Methyl acetate	-0.0287	-0.0528	72.6	46.2	39.5	11	xyTp	313.2	313.2	735.0	7974.6	Ohgaki K., Katayama T., J. Chem. Eng. Data, 1975, 20, 3, 264–267
Carbon dioxide	Methylcyclo-hexane	0.1782	0.1492	225.9	241.6	264.4	11	xTp	302.2	302.2	577.0	6484.0	Nasrifar Kh. et al., Fluid Phase Equilibria, 2003, 204, 1–14
Carbon dioxide	p-Cresol	0.0803	0.0709	48.4	35.9	98.3	10	xyTp	423.2	423.2	2000.0	20000.0	Lee M. *et al.,* Fluid Phase Equilibria, 1999, 162, 211–224
Carbon dioxide	Phenol	0.0731	0.0666	19.3	16.7	237.6	5	xyTp	423.2	423.2	1013.3	5066.3	Yau J., Tsal F., J. Chem. End. Data, 1992, 37, 141–143
Carbon dioxide	Toluene	0.1108	0.1171	37.0	24.9	59.3	8	xyTp	477.0	477.0	1179.0	15223.6	Ng H.-J., Robinson D. B., J. Chem. Eng. Data, 1978, 23, 325–327
Carbon dioxide	Water	-0.0660	0.1245	102.9	44.1	24.0	7	xyTp	598.2	598.2	14999.8	43499.3	Takenouchi S., Kennedy G. C., Am. J. Sci., 1964, 262, 1055
Cyclohexane	Acetic acid	-0.0118	0.0941	39.0	46.5	42.4	12	xyTp	352.0	391.3	101.3	101.3	Baradarajan A, Satyanarayana M., Indian J. Technol., 1967, 5, 264
Cyclohexane	Cyclohexene	0.0028	0.0066	1.8	3.9	1.7	21	xyTp	353.8	355.9	101.3	101.3	Marrufo B. *et al.,* Fluid Phase Equilibria, 2009, 279, 1, 11–16
Cyclohexane	Guaiacol	-0.0238	-0.0280	22.2	19.5	212.5	11	xTp	323.2	323.2	32.6	36.2	Ksiazczak A., Kosinski J. J., J. Chem. Eng. Data, 1991, 36, 351
Cyclohexane	m-Cresol	-0.0001	-0.0030	99.0	107.2	86.5	9	xTp	313.2	313.2	21.7	24.5	Azimi-Pour H., Rev. Inst. Fr. Pet. Ann. Combust. Liq., 1960, 15, 1–52
Cyclohexane	Methylcyclo-hexane	0.0024	0.0025	22.5	22.1	25.2	8	xyTp	356.8	371.5	101.3	101.3	Richards A. R., Hargreaves E., Ind. Eng. Chem., 1944, 36, 9, 805–808
Cyclohexane	Toluene	0.0053	0.0113	5.5	5.2	11.2	15	xyTp	420.1	451.1	506.6	506.6	Chen G. *et al.,* J. Chem. Ind. Eng. (China), 1994, 45, 94
Ethanol	1,2-Ethanediol	-0.0135	-0.0049	17.6	28.2	46.1	11	xyTp	351.5	470.5	101.3	101.3	Li J. *et al.,* Fluid Phase Equilibrium, 2000, 169, 75–84

1st comp	2nd comp	PC-SAFT k_ij	SRK k_ij	PC-SAFT error *	SRK error*	PSRK error *	# points	Quality	T ange (K)		p range (kPa)		Source
Ethanol	Acetic acid	-0.0794	-0.0559	13.3	20.3	74.6	16	xyTp	353.4	388.2	101.3	101.3	Amer Amezaga S., Fernandez Biarge J., An. Quim. 1973, 69, 569
Ethanol	Benzene	0.0076	0.1212	31.5	14.0	14.8	11	xyTp	388.0	408.4	446.1	446.1	Ellis S. R. M., Clark M. B., Chem. Age India, 1961, 12, 377
Ethanol	Catechol	0.0067	0.0928	136.1	110.2	341.4	5	xTp	290.2	290.2	2.0	4.1	Weissenberger G. *et al.,* Monatsh. Chem., 1925, 46, 471
Ethanol	Cyclohexane	0.0460	0.1090	11.4	37.7	21.4	14	xyTp	347.4	350.5	101.3	101.3	Zhao J. *et al.,* Fluid Phase Equilibria, 2006, 242, 147–153
Ethanol	Guaiacol	-0.0954	-0.0129	75.3	58.5	84.9	6	xTp	290.2	290.2	1.0	3.6	Weissenberger G. *et al.,* Monatsh. Chem., 1925, 46, 471
Ethanol	Hydrogen	-6.0337	-14.6570	2624.1	2627.0	3343.0	8	xTp	498.2	498.2	33300.0	10290.0	Brunner E., Hueltenschmidt W., J. Chem. Thermodyn., 1990, 22, 73–84
Ethanol	m-Cresol	-0.1470	-0.0782	101.8	66.5	64.6	11	xTp	351.5	474.7	101.3	101.3	Piatti L., Z. Phys. Chem., 1930, 152, 36
Ethanol	Methylcyclo-hexane	0.0496	0.1170	9.6	57.3	9.4	22	xyTp	345.3	368.8	101.3	101.3	Sànchez-Russinyol M. d. C. *et al.,* J. Chem. Eng. Data, 2004, 49, 1258–1262
Ethanol	o-Cresol	-0.1090	-0.0143	64.8	13.2	103.9	14	xTp	350.1	461.9	95.8	95.8	Prasad T. E. V. *et al.,* Fluid Phase Equilib., 2006, 244, 86–98
Ethanol	p-Cresol	-0.1504	-0.0729	75.1	57.5	74.3	12	xTp	350.0	472.9	95.2	95.2	Prasad T. E. V. *et al.,* Fluid Phase Equilib., 2006, 244, 86–98
Ethanol	Phenol	-0.0977	-0.0253	58.3	27.9	83.2	6	xTp	288.2	288.2	0.8	3.2	Weissenberger G. *et al.,* Monatsh. Chem., 1924, 45, 425
Ethanol	Toluene	0.0235	0.0998	11.2	26.9	17.1	12	xyTp	349.7	383.7	101.3	101.3	Aleykutty T. K., Srinivasan D., Indian J., Technol., 1975, 13, 345
Ethanol	Water	-0.0254	-0.1045	9.7	15.0	12.9	19	xyTp	409.9	435.3	666.9	666.9	Otsuki H., Williams F. C., Chem. Eng. Data Series, 1953, 49, 55

1st comp	2nd comp	PC-SAFT k_ij	SRK k_ij	PC-SAFT error *	SRK error*	PSRK error *	# points	Quality	T ange (K)		p range (kPa)		Source
Ethyl acetate	Acetic acid	-0.2068	-0.0351	75.2	40.4	66.3	26	xyTp	350.6	388.9	101.3	101.3	Calvar N. *et al.,* Fluid Phase Equilibria, 2005, 235, 215–222
Ethyl acetate	Benzene	-0.0407	0.0043	6.3	1.2	2.8	19	xyTp	350.3	353.2	101.3	101.3	Carr A. D., Kropholler H. W., J. Chem. Eng. Data, 1962, 7, 26
Ethyl acetate	Cyclohexane	0.0032	0.0752	4.2	3.9	2.6	13	xyTp	344.8	350.7	101.3	101.3	Chao K. C., Thesis, Wisconsin, 1956
Ethyl acetate	Ethanol	-0.0716	0.0340	17.7	5.9	7.3	24	xyTp	345.2	350.8	101.3	101.3	Calvar N. *et al.,* Fluid Phase Equilibria, 2005, 235, 215–222
Ethyl acetate	m-Cresol	0.0083	0.1026	227.6	218.3	1073.6	7	xTp	288.2	288.2	1.2	4.9	Weissenberger G. *et al.,* Monatsh. Chem., 1925, 46, 1
Ethyl acetate	o-Cresol	0.0180	0.1024	272.8	274.4	1018.3	7	xTp	288.2	288.2	2.5	5.8	Weissenberger G. *et al.,* Monatsh. Chem., 1925, 46, 1
Ethyl acetate	p-Cresol	0.0079	0.0745	219.0	243.2	1112.4	7	xTp	288.2	288.2	1.2	5.5	Weissenberger G. *et al.,* Monatsh. Chem., 1925, 46, 1
Ethyl acetate	Toluene	-0.0335	0.0163	10.0	4.9	6.1	18	xyTp	350.8	382.0	101.3	101.3	Carr A. D., Kropholler H. W., J. Chem. Eng. Data, 1962, 7, 26
Formic acid	Acetic acid	0.0230	0.0186	14.7	21.0	27.2	17	xyTp	374.0	391.3	101.3	101.3	Alpert N., Elving P. J., Ind. Eng. Chem., 1949, 41, 2864
Formic acid	Benzene	0.0069	0.1521	579.4	110.9	103.6	12	xyTp	333.1	333.1	25.4	69.1	Vrevsky M. S. *et al.*, Z. Phys. Chem., 1928, 133, 377
Formic acid	Carbon dioxide	-0.0026	0.2248	433.8	452.0	509.5	9	xTp	393.2	393.2	9180.0	23260.0	Byun H.-S. *et al.*, Ind. Eng. Chem. Res., 2000, 39, 4580–4587
Formic acid	Methyl acetate	-0.1624	-0.0519	61.9	27.6	154.7	23	xyTp	330.3	373.4	101.0	101.0	Fu H. *et al.,* Zhejiang Daxue Xuebao, 1987, 21, 52
Formic acid	Toluene	0.0205	0.1375	528.8	34.5	57.7	19	xyTp	298.2	298.2	3.8	8.7	Lakhanpal M. L. *et al.,* Indian J. Chem., 1975, 13, 1309

1st comp	2nd comp	PC-SAFT k_ij	SRK k_ij	PC-SAFT error *	SRK error*	PSRK error *	# points	Quality	T ange (K)		p range (kPa)		Source
Hydrogen	Acetic acid	-0.0422	-1.2393	149.9	155.3	634.6	17	xyTp	343.2	343.2	270.0	7400.0	Jonasson A. *et al.,* Fluid Phase Equilibria, 1998, 152, 67–94
Hydrogen	Benzene	-0.0907	2.6791	5.9	25.4	59.1	3	xyTp	533.2	533.2	4158.4	11470.0	Connolly J. F., The Journal of Chemical Physics, 1962, 36, 11, 2897–2904
Hydrogen	Carbon dioxide	-0.6452	0.1033	12.1	30.6	48.5	6	xyTp	289.9	289.9	6575.9	17528.9	Barrick P. L. *et al.,* Tech. Rep. Afml TR, 1966, 66–390
Hydrogen	Cyclohexane	0.0689	0.1462	178.7	180.2	237.4	11	xTp	303.2	303.2	887.0	4743.0	Tsuji T. *et al.,* Fluid Phase Equilibrium, 2005, 228–229, 499–503
Hydrogen	Cyclohexene	0.1334	0.5071	225.1	224.0	394.2	5	xTp	373.0	373.0	790.8	6996.1	Herskowitz M. *et al.,* J. Chem. Eng. Data, 1983, 28, 164–166
Hydrogen	Ethyl acetate	-0.1724	0.1719	181.9	181.8	264.1	4	xTp	291.0	291.0	1600.0	4600.0	Wainwright M. S. *et al.,* J. Chem. Eng. Data, 1987, 32, 22
Hydrogen	m-Cresol	-0.1822	3.9926	13.4	44.8	41.7	7	xyTp	662.0	662.0	2681.0	25370.0	Simnick J. J. *et al.,* J. Chem. Thermodyn., 1979, 11, 531–537
Hydrogen	Methane	-0.4740	-0.0182	36.4	36.3	30.2	9	xyTp	173.7	173.7	3566.6	10831.6	Sagara H. *et al.,* J. Chem. Eng. Data, 1972, 5, 4, 339–348
Hydrogen	Methanol	-0.3113	-1.4443	182.8	169.3	205.5	6	xyTp	308.2	308.2	2000.0	11000.0	Bezanehtak G. B. *et al.,* J. Chem Eng. Data 2002, 47, 161–168
Hydrogen	Methyl acetate	-0.1267	0.2912	146.4	139.3	201.1	4	xTp	291.0	291.0	1150.0	3110.0	Wainwright M. S. *et al.,* J. Chem. Eng. Data, 1987, 32, 22
Hydrogen	Methylcyclo-hexane	0.0665	0.1081	199.6	202.4	234.6	8	xTp	303.2	303.2	1236.0	4323.0	Tsuji T. *et al.,* Fluid Phase Equilibrium, 2005, 228–229, 499–503
Hydrogen	Toluene	0.1131	2.0756	22.4	133.2	90.4	5	xyTp	575.2	575.2	5015.6	25351.5	Simnick J. J. *et al.,* J. Chem. Eng. Data, 1978, 23, 4, 339–340
Hydrogen	Water	-0.3747	-2.2631	142.3	96.1	332.7	8	xTp	373.2	373.2	2533.1	101325.0	Krichevsky I. R., Kasarnovsky J. S., J. Am. Chem. Soc., 1935, 57, 11, 2168–2171

1st comp	2nd comp	PC-SAFT k_ij	SRK k_ij	PC-SAFT error *	SRK error*	PSRK error *	# points	Quality	T ange (K)		p range (kPa)		Source
m-Cresol	Carbon dioxide	0.0921	0.1058	135.2	120.5	125.3	5	xyTp	373.2	373.2	10200.0	30000.0	Pfohl O. *et al.,* Fluid Phase Equilibria, 1997, 141, 179–206
Methane	Acetic acid	0.2300	0.2256	128.1	130.7	158.0	13	xyTp	348.0	348.0	310.0	6250.0	Jonasson A. *et al.,* Fluid Phase Equilib., 1998, 152, 67–94
Methane	Benzene	0.0477	0.1231	12.2	26.9	47.6	6	xyTp	501.2	501.2	5123.9	14557.1	Lin H. M. *et al.,* J. Chem. Eng. Data, 1979, 24, 2, 146–149
Methane	Carbon dioxide	0.0602	0.1096	16.2	7.4	37.2	10	xyTp	270.0	270.0	3198.8	8432.3	Somait F. A., Kiday A. J., J. Chem. Eng. Data, 1978, 23, 4, 301
Methane	Cyclohexane	0.0549	0.0648	7.4	44.9	39.7	14	xyTp	444.3	444.3	827.4	20063.6	Reamer H. H. *et al.,* Ind. Eng. Chem. Data Ser., 1958, 3, 2, 240
Methane	Ethanol	0.0278	0.0344	5.9	26.0	74.0	5	xyTp	333.4	333.4	2594.0	10464.0	Suzuki K., Sue H., J. Chem. Eng. Data, 1990, 35, 63–66
Methane	m-Cresol	0.0237	0.3093	36.9	27.9	59.2	4	xyTp	663.4	663.4	5116.8	15249.2	Simnick J. J. *et al.,* Fluid Phase Equilibria, 1979, 3, 145–151
Methane	Methanol	0.0495	-0.0172	59.4	72.9	92.1	11	xyTp	330.0	330.0	1379.0	41368.5	Hong J. H., Malone P. V., Jett M. D., Kobayashi R., Fluid Phase Equilib., 1987, 38, 83–96
Methane	Methyl acetate	-0.0134	0.0852	95.8	91.0	517.0	6	xTp	298.2	298.2	101.3	7990.0	Xia S. *et al.,* CHEMICAL ENGINEERING(CHINA), 2005, 33, 4, 68
Methane	Methylcyclohexane	0.0907	0.1274	10.9	21.8	58.0	8	xyTp	473.2	473.2	3350.0	18020.0	Richon D. *et al.,* J. Chem. Eng. Data, 1991, 36, 104
Methane	Toluene	0.0494	0.1787	18.2	34.5	59.7	5	xyTp	543.2	543.2	3077.2	11510.3	Lin H. M. *et al.,* J. Chem. Eng. Data, 1979, 24, 2, 146–149
Methanol	1,2-Ethanediol	-0.0255	0.0045	70.7	84.1	83.0	14	xyTp	337.3	472.2	101.3	101.3	Baker T. H. *et al.,* J. Chem. Eng. Data, 1964, 9, 11
Methanol	Acetic acid	-0.0884	-0.0548	10.9	34.4	75.4	19	xyTp	338.9	387.9	101.3	101.3	Amer A. S., Fernandez B. J., An. Quim., 1973, 69, 569

1st comp	2nd comp	PC-SAFT k_ij	SRK k_ij	PC-SAFT error *	SRK error*	PSRK error *	# points	Quality	T ange (K)		p range (kPa)		Source
Methanol	Benzaldehyde	-0.0150	0.2438	269.4	301.3	849.3	6	xTp	293.2	293.2	7.5	12.3	Weissenberger G. *et al.,* Monatsh. Chem., 1925, 46, 47–56
Methanol	Benzene	0.0118	0.1242	14.2	33.0	7.3	10	xyTp	373.2	373.2	312.3	417.8	Butcher K. L., Medani M. S., J. Appl. Chem., 1968, 18, 100
Methanol	Cyclohexane	0.0575	0.1122	12.2	56.1	19.0	13	xyTp	327.4	353.9	101.3	101.3	Budantseva L. S. *et al.,* Zh. Fiz. Khim., 1975, 49, 260
Methanol	Cyclohexene	0.0416	0.1145	21.0	48.2	24.6	13	xyTp	329.2	356.2	101.3	101.3	Budantseva L. S. *et al.,* Zh. Fiz. Khim., 1975, 49, 260
Methanol	Ethanol	-0.0018	-0.0147	2.5	8.0	23.1	10	xyTp	393.2	393.2	427.8	620.5	Piatti L., Z. Phys. Chem., 1930, 152, 36
Methanol	Ethyl acetate	-0.1076	0.0208	49.8	18.7	6.0	19	xyTp	335.3	348.0	101.3	101.3	Akita K., Yoshida F., J. Chem. Eng. Data, 1963, 8, 484
Methanol	Guaiacol	-0.1919	0.0036	238.7	133.1	144.4	10	xyTp	338.0	478.0	101.3	101.3	Chalov N. V. *et al.*, Kh. Gidroliz. Lesokhim. Prom-st., 1955, 8, 11
Methanol	m-Cresol	-0.2647	-0.1377	184.1	132.3	214.3	14	xTp	336.3	473.8	95.5	95.5	Prasad T. E. V. *et al.,* J. Chem. Thermodyn., 2006, 38, 1696–1700
Methanol	Methylcyclo-hexane	0.0638	0.0928	15.2	123.1	19.9	30	xyTp	333.2	333.2	27.1	102.9	Scheller M. *et al.,* J. Prakt. Chem., 1969, 311, 974
Methanol	o-Cresol	-0.2159	-0.0379	111.3	39.0	46.1	9	xyTp	337.9	463.9	101.3	101.3	Chalov N. V. *et al.,* Gidroliz. Lesokhim. Prom-st., 1955, 8, 11
Methanol	p-Cresol	-0.2901	-0.1489	211.2	176.5	274.8	12	xTp	336.2	472.9	95.2	95.2	Prasad T. E. V. *et al.,* Fluid Phase Equilib., 2006, 244, 86–98
Methanol	Toluene	0.0151	0.1091	27.9	61.9	17.2	13	xyTp	336.8	383.8	101.3	101.3	Budantseva L. S. *et al.,* Zh. Fiz. Khim., 1975, 49, 1844
Methanol	Water	-0.0338	-0.0753	11.4	13.3	14.0	22	xyTp	386.0	421.2	506.6	506.6	Hirata M. *et al.,* Computer aided data Book of vapor-liquid equilibria, Elsevier 1975

1st comp	2nd comp	PC-SAFT k_ij	SRK k_ij	PC-SAFT error *	SRK error*	PSRK error *	# points	Quality	T ange (K)		p range (kPa)		Source
Methyl acetate	Acetic acid	-0.1982	-0.0113	51.6	41.5	50.2	14	xyTp	332.6	377.7	101.3	101.3	Balashov M. I., Serafimov L. A., Khim. Khim. Technol., 1966, 9, 885
Methyl acetate	Benzene	-0.0327	0.0321	20.1	20.4	25.5	11	xTp	363.2	363.2	135.3	299.6	Schmidt G. C., Z. Phys. Chem., 1926, 121, 221
Methyl acetate	Cyclohexane	0.0105	0.1009	7.3	10.4	9.6	14	xyTp	328.7	347.5	101.3	101.3	Nagata I., J. Chem. Eng. Data, 1962, 7, 461
Methyl acetate	Ethanol	-0.0961	0.0276	31.1	36.2	42.4	7	xTp	333.2	333.2	47.0	99.7	Perelygin V. M., Suntsov Yu. K., Izv. Vyssh. Uchebn. Zaved. Pishch. Tekhnol., 1974, 3, 94
Methyl acetate	Ethyl acetate	0.0142	0.0209	18.4	16.8	27.0	11	xTp	353.2	353.2	122.4	211.7	Schmidt G. C., Z. Phys. Chem., 1926, 121, 221
Methyl acetate	Methanol	-0.1087	0.0315	37.6	10.3	8.9	20	xyTp	326.9	337.9	101.3	101.3	Mato F., Cepeda E., An. Quim. Ser. A., 1984, 80, 338
Methyl acetate	Toluene	-0.0306	0.0384	7.1	8.9	15.1	24	xyTp	331.0	379.7	101.0	101.0	Wisniak J., Tamir A., J. Chem. Eng. Data, 1989, 34, 16–19
Methyl acetate	Water	-0.1129	-0.1988	38.1	78.5	52.5	8	xyTp	353.2	353.2	83.1	218.4	Perelygin V. M., Volkov A. G., Izv. Vyssh. Ucheb. Zaved. Pishch. Tekhnol., 1970, 124
Methylcyclo-hexane	Phenol	0.0367	0.0125	105.2	113.4	111.4	9	xyTp	374.3	423.2	101.3	101.3	Drickamer H. G. *et al.*, Trans. Am. Inst. Chem. Eng., 1945, 41, 555
Methylcyclo-hexane	Toluene	0.0114	0.0291	2.1	22.8	1.5	29	xyTp	348.2	348.2	32.6	45.7	Diaz Pena M., Compostizo A., J. Chem. Eng. Data, 1983, 28, 30
o-Cresol	1,2-Ethanediol	-0.0351	0.0087	18.4	8.8	16.4	13	xyTp	462.0	468.1	100.0	100.0	Brusset H. *et al.,* Chim. Ind. Genie Chim., 1968, 99, 207
o-Cresol	Carbon dioxide	0.0638	0.0574	172.4	146.2	200.3	5	xyTp	373.2	373.2	10400.0	26000.0	Pfohl O. *et al.,* Fluid Phase Equilibria, 1997, 141, 179–206
o-Cresol	m-Cresol	0.0005	-0.0022	2.4	2.9	2.7	21	xTp	464.4	475.5	101.0	101.0	Fox J. J., Barker M. F., J. Soc. Chem. Ind., London, 1917, 36, 842

1st comp	2nd comp	PC-SAFT k_ij	SRK k_ij	PC-SAFT error *	SRK error*	PSRK error *	# points	Quality	T ange (K)		p range (kPa)		Source
o-Cresol	p-Cresol	-0.0061	-0.0106	8.2	7.5	16.3	22	xTp	462.4	473.2	96.4	96.4	Selvam S. P. *et al.,* Fluid Phase Equilibria, 1992, 78, 261–267
p-Cresol	Catechol	0.0231	0.0044	32.7	39.1	14.1	11	xyTp	453.2	453.2	15.5	55.7	Hwang S. *et al.,* Fluid Phase Equilibria, 2000, 172, 183–196
Phenol	1,2-Ethanediol	-0.0518	-0.0117	15.5	11.5	10.4	12	xyTp	454.7	468.5	100.0	100.0	Brusset H. *et al.,* Chim. Ind. Genie chim., 1968, 99, 207
Phenol	m-Cresol	-0.0002	-0.0035	10.8	16.5	12.8	5	xyTp	397.2	407.6	13.3	13.3	Vostrikova V. N. *et al.,* Khim. Prom., 1978, 5, 345
Phenol	Methanol	-0.2099	-0.0447	111.1	31.0	53.0	21	xyTp	337.9	455.0	101.3	101.3	Hu W. *et al.,* Fluid Phase Equilibria, 2004, 219, 265–268
Phenol	o-Cresol	0.0009	-0.0005	3.0	4.0	2.6	22	xTp	453.2	462.4	96.4	96.4	Selvam S. P. *et al.,* Fluid Phase Equilibria, 1992, 78, 261–267
Phenol	p-Cresol	-0.0050	-0.0096	7.1	6.8	16.8	21	xTp	453.2	473.2	96.4	96.4	Selvam S. P. *et al.,* Fluid Phase Equilibria, 1992, 78, 261–267
Toluene	1,2-Ethanediol	0.0051	0.0030	244.2	190.8	271.7	61	xTp	323.2	323.2	0.3	12.5	Witting R. *et al.,* Ind. Eng. Chem. Res., 2001, 40, 5831–5838
Toluene	Acetic acid	-0.0431	0.0878	42.0	21.4	26.5	22	xyTp	377.6	391.2	101.3	101.3	Haughton C. O., Chem Eng. Sci., 1961, 16, 82
Toluene	Benzaldehyde	0.0084	0.0160	44.8	46.3	52.6	13	xyTp	383.8	452.2	101.3	101.3	Rao K. V. *et al.,* J. Chem. Eng. Data, 1981, 26, 413
Toluene	m-Cresol	0.0210	0.0320	23.8	31.8	24.4	13	xyTp	363.2	363.2	11.5	53.3	Sartakova O. Yu. *et al.,* Zh. Prikl. Khim. (S.-Petersburg), 1996, 69, 1077
Toluene	o-Cresol	0.0132	0.0121	26.1	27.7	29.2	8	xyTp	363.2	363.2	12.4	48.0	Klauck M. *et al.,* Ind. Eng. Chem. Res., 2009, 47, 15, 5119–5126
Toluene	p-Cresol	0.0191	0.0301	18.0	20.2	14.7	19	xyTp	383.9	475.4	101.3	101.3	Marks D. E., Wingard R. E., J. Chem. Eng. Data, 1960, 5, 416

1st comp	2nd comp	PC-SAFT k_ij	SRK k_ij	PC-SAFT error *	SRK error*	PSRK error *	# points	Quality	T ange (K)		p range (kPa)		Source
Toluene	Phenol	0.0270	0.0189	15.2	16.3	21.3	21	xyTp	383.7	445.9	101.3	101.3	Drickamer H. G. *et al.,* Trans. Am. Inst. Chem. Eng., 1945, 41, 555
Toluene	Water	-0.1150	-0.2695	1588.5	1580.4	2054.3	4	xTp	298.0	473.0	5000.0	5000.0	Mathis J. *et al.,* J. Chem. Eng. Data, 2004, 49, 1269–1272
Water	1,2-Ethanediol	0.0087	-0.0095	22.6	28.4	70.6	7	xTp	371.7	469.7	97.3	97.3	Kumar P. *et al.,* Indian Chem. Eng., 1982, 24, 14
Water	Acetic acid	-0.1068	-0.1565	6.6	16.5	44.0	17	xyTp	403.6	426.8	273.7	273.7	Othmer D. F. *et al.,* Ind. Eng. Chem., 1952, 44, 1864
Water	Cyclohexane	0.3693	0.2623	387.2	458.9	164.8	7	xTp	627.0	651.1	80000.0	80000.0	Broellos K. *et al.,* Ber. Bunsen-Ges. Phys. Chem., 1970, 74, 682–686
Water	Ethyl acetate	-0.0986	-0.2540	7.6	7.9	47.6	6	xyTp	369.2	373.2	101.3	101.3	Anoshin I. M., Salnikov A. P., Izv. Vyssh. Uchebn. Zaved., Pishch. Tekhnol., 1976, 4, 145
Water	Formic acid	-0.1520	-0.2662	16.4	17.2	73.6	14	xyTp	407.1	417.1	298.2	298.2	Gilburd M. M. *et al.,* Zh. Prikl. Khim., 1984, 57, 915
Water	m-Cresol	-0.0605	-0.0656	53.8	32.6	32.3	7	xyTp	363.2	363.2	35.6	70.3	Klauck M. *et al.,* Ind. Eng. Chem. Res., 2009, 47, 15, 5119–5126
Water	Methane	0.0159	-0.3001	69.8	178.4	236.2	6	xyTp	318.1	318.1	1003.0	34610.0	Chapoy A. *et al.,* Fluid Phase Equilibria, 2003, 214, 101–117
Water	o-Cresol	-0.0248	-0.0633	63.4	134.6	98.7	9	xyTp	372.2	443.7	101.3	101.3	Brusset H., Gaynes J., Comp. Rend., 1953, 236, 1563
Water	Phenol	-0.0348	-0.0254	61.0	143.0	118.6	18	xyTp	372.8	455.2	101.3	101.3	Brusset H., Gaynes J., Comp. Rend., 1953, 236, 1563

* The relative error is provided by Aspen Plus software.

Appendix B: Langmuir-Hinshelwood reaction rate by Horiuti-Polanyi mechanism for Reaction 4.3.1

$H_2 + 2* \rightleftharpoons 2H*$	$r_{A1} = k_{A1} C_{H_2} C_v^2 - k_{-A1} C_{H*}^2 = 0$
$A + H* \rightleftharpoons A \cdot H*$	$r_{A2} = k_{A2} C_A C_{H*} - k_{-A2} C_{A \cdot H*} = 0$
$A \cdot H* \rightharpoonup A'*$	$r_{S1} = k_{S1} C_{A \cdot H*}$
$A'* \rightleftharpoons A* + *$	$r_{D1} = k_{D1} C_{A'*} - k_{-D1} C_{A'} C_v = 0$
$A' + H* \rightleftharpoons A' \cdot H*$	$r_{A3} = k_{A3} C_{A'} C_{H*} - k_{-A3} C_{A \cdot H*} = 0$
$A' \cdot H* \rightleftharpoons A''*$	$r_{S2} = k_{S2} C_{A' \cdot H*} - k_{-S2} C_{A''*} = 0$
$A''* \rightleftharpoons A'' + *$	$r_{D2} = k_{D2} C_{A''*} - k_{-D2} C_{A''} C_v = 0$
$A'' + H* \rightleftharpoons A'' \cdot H*$	$r_{A4} = k_{A4} C_{A''} C_{H*} - k_{-A4} C_{A'' \cdot H*} = 0$
$A'' \cdot H* \rightleftharpoons A'''*$	$r_{S3} = k_{S3} C_{A'' \cdot H*} - k_{-S3} C_{A'''*} = 0$
$A'''* \rightleftharpoons A''' + *$	$r_{D3} = k_{D3} C_{A'''*} - k_{-D3} C_{A'''} C_v = 0$
$A''' + H* \rightleftharpoons A''' \cdot H*$	$r_{A5} = k_{A5} C_{A'''} C_{H*} - k_{-A5} C_{A''' \cdot H*} = 0$
$A''' \cdot H* \rightleftharpoons B*$	$r_{S4} = k_{S4} C_{A''' \cdot H*} - k_{-S4} C_{B*} = 0$
$B* \rightleftharpoons B + *$	$r_{D4} = k_{D4} C_{B*} - k_{-D4} C_B C_v = 0$

$$C_t = C_v + C_{H*} + C_{A \cdot H*} + C_{A'*} + C_{A' \cdot H*} + C_{A''*} + C_{A'' \cdot H*} + C_{A'''*} + C_{A''' \cdot H*} + C_{B*}$$

$$C_{H*} = C_v (K_{A1} C_{H_2})^{0.5}$$

$$C_{A \cdot H*} = K_{A2} C_A C_{H*} = K_{A2} C_A C_v (K_{A1} C_{H_2})^{0.5}$$

$$C_{A'*} = K_{D1}^{-1} C_{A'} C_v$$

$$C_{A' \cdot H*} = K_{A3} C_{A'} C_{H*} = K_{A3} C_{A'} C_v (K_{A1} C_{H_2})^{0.5}$$

$$C_{A''*} = K_{S2} C_{A' \cdot H*} = K_{S2} K_{A3} C_{A'} C_v (K_{A1} C_{H_2})^{0.5}$$

$$C_{A''*} = K_{D2}^{-1} C_{A''} C_v$$

$$C_{A''\cdot H\bullet} = K_{A4} C''_A C_{H*} = K_{A4} C''_A C_v \left(K_{A1} C_{H_2}\right)^{0.5}$$

$$C_{A'''\bullet} = K_{S3} C_{A''\cdot H\bullet} = K_{S3} K_{A4} C''_A C_v \left(K_{A1} C_{H_2}\right)^{0.5}$$

$$C_{A'''\bullet} = K_{D3}^{-1} C_{A'''} C_v$$

$$C_{A'''\cdot H\bullet} = K_{A5} C_{A'''} C_{H\bullet} = K_{A5} C_{A'''} C_v \left(K_{A1} C_{H_2}\right)^{0.5}$$

$$C_{B\bullet} = K_{S4} C_{A'''\cdot H\bullet} = K_{S4} K_{A5} C_{A'''} C_v \left(K_{A1} C_{H_2}\right)^{0.5}$$

$$C_{B\bullet} = K_{D4}^{-1} C_B C_v$$

$$C_t = C_v + C_{H\bullet} + C_{A\cdot H\bullet} + C_{A'\bullet} + C_{A'\cdot H\bullet} + C_{A''\bullet} + C_{A''\cdot H\bullet} + C_{A'''\bullet} + C_{A'''\cdot H\bullet} + C_{B\bullet}$$

$$r_{S1} = k_{S1} K_{A2} C_A C_v \left(K_{A1} C_{H_2}\right)^{0.5}$$

$$r_{S1} = {k_{S1} K_{A2} C_A C_t \left(K_{A1} C_{H_2}\right)^{0.5}} \Big/ {1 + \left(K_{A1} C_{H_2}\right)^{0.5} + K_{A2} C_A \left(K_{A1} C_{H_2}\right)^{0.5} + K_{D1}^{-1} C_{A'} + K_{A2} C_{A'} \left(K_{A1} C_{H_2}\right)^{0.5} + K_{D2}^{-1} C_{A''} + K_{A4} C''_A \left(K_{A1} C_{H_2}\right)^{0}}$$

$$r_{S1} = \frac{k_{tot} C_A \left(C_{H_2}\right)^{0.5}}{1 + K_A C_A \left(C_{H_2}\right)^{0.5} + K_{D4}^{-1} C_B}$$

Equivalently for Reaction 4.3.2;

$$r_{S1} = \frac{k_{tot} C_B \left(C_{H_2}\right)^{0.5}}{1 + K_A C_B \left(C_{H_2}\right)^{0.5} + K_{D2}^{-1} C_C}$$

Equivalently for Reaction 4.3.3;

$$r_{S1} = \frac{k_{tot} C_C \left(C_{H_2}\right)^{0.5}}{1 + K_A C_C \left(C_{H_2}\right)^{0.5} + K_{D4}^{-1} C_D}$$

Equivalently for Reaction 4.3.4;

$$r_{S1} = \frac{k_{tot} C_D \left(C_{H_2}\right)^{0.5}}{1 + K_A C_D \left(C_{H_2}\right)^{0.5} + K_{D2}^{-1} C_D}$$

Appendix C: The experimental data used for the optimization the kinetic parameters for guaiacol HDO reaction

Table C.1. Experimental data at 175°C (concentration in mol/dm^3).

Time (min)	C_GUA	C_prod-ucts	C_methoxycyclo-hexanone	C_methyl-1,2-cyclohexanediol	C_cyclohex-anol	C_cyclohex-ane
0	0,136231	0	0	0	0	0
10	0,116574	0.007551	0,00136	0,006191	0	0
20	0,10954	0.015373	0,002688	0,012024	0,000661	0
40	0,100126	0.030156	0,004545	0,024337	0,001274	0
60	0,085517	0.045089	0,00536	0,037919	0,00181	0
80	0,077904	0.05516	0,005275	0,047599	0,002286	0
100	0,064417	0.062786	0,004356	0,055847	0,002583	0

Table C.2. Experimental data at 175°C (concentration in mol/dm^3).

Time (min)	C_GUA	C_prod-ucts	C_methoxycyclo-hexanone	C_methyl-1,2-cyclohexanediol	C_cyclohex-anol	C_cyclohex-ane
0	0,137423	0	0	0	0	0
10	0,114452	0.011145	0,002016	0,008671	0,000458	0
20	0,099087	0.025831	0,003756	0,02117	0,000905	0
30	0,084072	0.04356	0,004636	0,037464	0,00146	0

40	0,071691	0.05552	0,004532	0,049166	0,001821	0
60	0,059871	0.067342	0,003865	0,061272	0,002206	0
120	0,04793	0.076065	0,00286	0,070686	0,00252	0

Table C.3. Experimental data at 200°C (concentration in mol/dm^3).

Time (min)	C_GUA	C_prod-ucts	C_methoxycyclo-hexanone	C_methyl-1,2-cyclohexanediol	C_cyclohex-anol	C_cyclohex-ane
0	0,138445	0	0	0	0	0
15	0,05419	0.059921	0,029761	0,028919	0,001241	0
2	0,009762	0.116237	0,022016	0,08916	0,004469	0,000592
70	0,002077	0.125068	0,009634	0,108953	0,005851	0,00063
80	0,000905	0.129047	0,006192	0,115684	0,006489	0,000682
90	0,000291	0.124317	0,003309	0,113891	0,006456	0,000661
100	0	0.120736	0,001866	0,111595	0,006605	0,000669

Table C.4. Experimental data at 200°C (concentration in mol/dm^3).

Time (min)	C_GUA	C_prod-ucts	C_methoxycyclo-hexanone	C_methyl-1,2-cyclohexanediol	C_cyclohex-anol	C_cyclohex-ane
0	0,137839	0	0	0	0	0
10	0,105121	0.015709	0,007169	0,007988	0,000552	0
20	0,045043	0.079219	0,028574	0,047522	0,002631	0,000491
30	0,020898	0.105227	0,029433	0,071918	0,003877	0
40	0,014328	0.114213	0,026316	0,082968	0,004505	0,000424
60	0,004557	0.126265	0,014652	0,105186	0,005976	0,000451
120	0,003229	0.126745	0,010149	0,109144	0,006752	0,000701

Table C.5. Experimental data at 220°C (concentration in mol/dm^3).

Time (min)	C_GUA	C_prod-ucts	C_methoxycyclo-hexanone	C_methyl-1,2-cyclohexanediol	C_cyclohex-anol	C_cyclohex-ane
0	0,171643	0	0	0	0	0
15	0,089447	0.033077	0,018776	0,011973	0,002328	0
50	0,007961	0.159852	0,019333	0,120499	0,02002	0
70	0,00119	0.112843	0,002384	0,088023	0,022436	0

80	0,000667	0.114472	0,00138	0,088472	0,024621	0
90	0,000377	0.114549	0,000771	0,08803	0,025748	0
100	0	0.113786	0	0,086324	0,027461	0

Table C. 6. Experimental data at 220°C (concentration in mol/dm^3).

Time (min)	C_GUA	C_prod-ucts	C_methoxycyclo-hexanone	C_methyl-1,2-cyclohexanediol	C_cyclohex-anol	C_cyclohex-ane
0	0,171535	0	0	0	0	0
10	0,130799	0.00345	0	0,002668	0,000782	0
20	0,04551	0.066965	0,031899	0,027842	0,006561	0,000663
30	0,020136	0.126469	0,032871	0,082743	0,010034	0,00082
40	0,012019	0.11907	0,021051	0,085335	0,011811	0,000873
60	0,00395	0.118973	0,010066	0,093366	0,014611	0,000931
120	0,001544	0.119757	0,002682	0,094303	0,021608	0,001164

Appendix D: Rate differences of the guaiacol HDO reaction

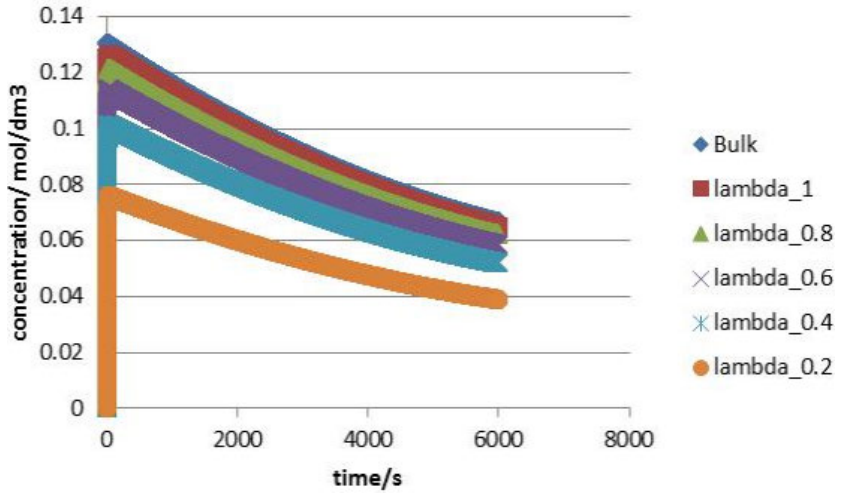

Figure D1. Concentration profile of guaiacol inside the catalyst caused by diffusion only for experiment 1 (Table C1).

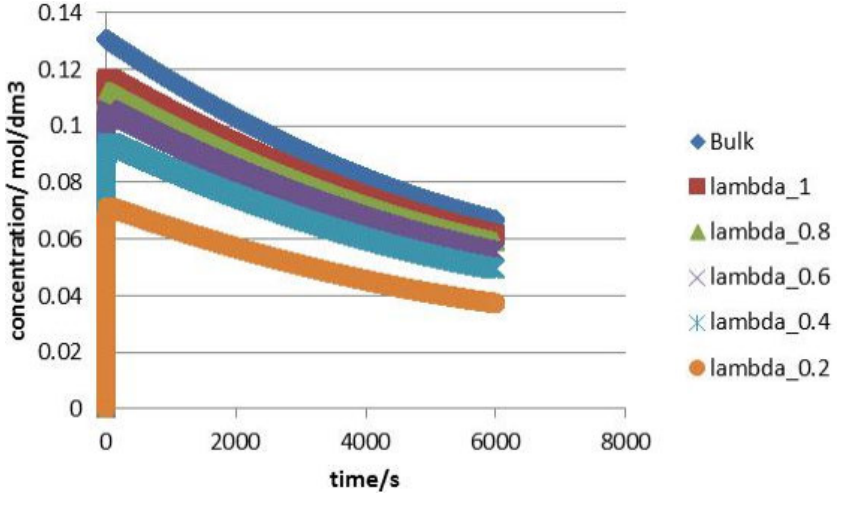

Figure D2. Concentration profile of guaiacol calculated with power law reaction rate equations, diffusion and reaction to methoxycyclohexane is taken into account. Experiment 1 (Table C1).

Appendix E: Input factors of the bark biorefinery sensitivity analysis

Code	Type	Factor description		Values		
				Nom.	Min	Max.
x2	Const.	Feed	Reference heating value (MW)	12.7		
x3	Var.	Feed	Fresh water to leaching (kg/kg dry bark)	10	5	20
x4	Var.	Hydrolysis and fermentation	Hydrolysis consistency	0.2	0.1	0.3
x5	Var.	Evaporation	Steam consumption kg steam/kg water evaporated	0.3	0.15	1
x6	Var.	Stilbene recovery	Separation efficiency	0.8	0	0.8
x7	Var.	Degree of water cycle closure		0.999		0.999
x8	Par.	Leaching	Leaching yields, Conversion of [Lignin]	0.275	0.2	0.3
x9	Par.	Leaching	Leaching yields, Conversion of [C5 Poly]	0.15	0.1	0.2
x10	Par.	Leaching	Leaching yields, Conversion of [C6 Poly]	0.25	0.2	0.3
x11	Par.	Leaching	Leaching yields, Conversion of [Tannin]	0.2	0.15	0.25
x12	Par.	Leaching	Leaching yields, Conversion of [Stilbene]	0.12	0.05	0.15
x13	Par.	Leaching	Leaching yields, Conversion of [Lipofilics]	0.6		
x14	Par.	Leaching	Leaching yields, Conversion of [Silica(s)]	0.2	0.15	0.3
x15	Par.	Leaching	Consistency of leached bark after dewatering	0.25	0.2	0.3
x16	Par.	Leaching	Consistency of bark after bark press	0.38	0.25	0.5
x17	Par.	Stilbene recovery	Purity	0.5	0.3	0.7
x18	Par.	Stilbene recovery	Dry content	0.2	0.1	0.4
x19	Par.	Tannin recovery	Conversion of polysugars to permeate-size molecules	0.01	0	0.75
x20	Par.	Tannin recovery	Volume reduction in ultrafiltration	3	1	5
x21	Par.	Tannin recovery	Ultrafiltration retentions, Tannin, lignin and stilbene	0.98	0.8	0.99

Code	Type	Factor description		Values		
				Nom.	Min	Max.
x22	Par.	Tannin recovery	Ultrafiltration retentions, Lipofilics	0.99	0.5	0.9
x23	Par.	Tannin recovery	Ultrafiltration retentions, Polysugars	0.7	0.5	0.8
x24	Par.	Hydrolysis and fermentation	Hydrolysis yield	0.7	0.6	0.8
x25	Par.	Hydrolysis and fermentation	Fermentation conversions, Conversion of [Glucose(aq)]	0.7	0.6	0.9
x26	Par.	Hydrolysis and fermentation	Fermentation conversions, Conversion of [Xylose(aq)]	0.01	0.01	0.9
x27	Par.	Distillation	Dry content after evaporation	0.3	0.2	0.5
x28	Par.	Unit prices	Fuel (€/MWh)	20	4	30
x29	Par.	Unit prices	Steam (€/ton)	12	3	16
x30	Par.	Unit prices	Tannin (€/ton)	300	200	1000
x31	Par.	Unit prices	Stilbene (€/ton)	3000	2000	10000
x32	Par.	Unit prices	Ethanol (€/ton)	500	400	800
x33	Par.	Unit prices	Enzyme (€/ton EtOH)	80	40	120
x34	Par.	Unit prices	COD in effluent (€/ton)	400	50	500

Series title and number
VTT Technology 67

Title	**BIOSCEN** **Modelling Biorefinery Scenarios**
Author(s)	Juha Leppävuori & Pertti Koukkari (eds.)
Abstract	Biorefining unit processes were studied in terms of their physicochemical modelling properties and parameters, complemented by surrogate meta-modelling and extensive sensitivity analysis. The unit processes included flash condensation of the fast biomass pyrolysis, thermal and enzyme catalysed hydrolysis of lignocellulosic biomass and fermentation of the hydrolysis product to bioethanol and biobutanol. In addition, a flowsheet based mass and energy balance was developed for the bark biorefinery, the key factors of which were then assessed by using the elementary efficiency method for sensitivity analysis. With the results from hydrolysis and fermentation models, a comparison of the greenhouse gas emissions from barley straw based ethanol and butanol was performed. The bark-based biorefinery producing mainly tannin and ethanol proved to be economically challenging. The results also show that biobutanol is in general more demanding (in terms of greenhouse gas emissions) than the production of bioethanol.
ISBN, ISSN	ISBN 978-951-38-7902-0 (URL: http://www.vtt.fi/publications/index.jsp) ISSN 2242-122X (URL: http://www.vtt.fi/publications/index.jsp)
Date	December 2012
Language	English
Pages	159 p. + app. 17 p.
Name of the project	BioScen
Commissioned by	
Keywords	Biorefinery, material properties, modelling, parameter optimisation, pyrolysis oil condensation, lignocellulosic biomass, hydrolysis, fermentation, greenhouse gas emissions
Publisher	VTT Technical Research Centre of Finland P.O. Box 1000, FI-02044 VTT, Finland, Tel. 020 722 111

Julkaisun sarja ja numero
VTT Technology 67

Nimeke	**BIOSCEN** **Biojalostamoprosessien mallinnus**
Tekijä(t)	Juha Leppävuori & Pertti Koukkari (toim.)
Tiivistelmä	Joukkoa biojalostamon yksikköprosesseja tutkittiin simulointimallien avulla. Tutkimuksessa tarkasteltiin tyypillisten biojalostamotuotteiden fysikaalis-kemiallisten ainearvojen saatavuutta ja näiden soveltamista termodynaamisissa ja kemialliseen reaktoritekniikkaan perustuvissa laskentamalleissa. Parametrien luotettavuutta testattiin metamallinnuksen ja herkkyysanalyysien avulla. Tutkitut yksikköprosessit olivat pyrolyysiöljyn talteenotto kondensoimalla, terminen ja entsyymikatalysoitu lignoselluloosan hydrolyysi sekä edelleen hydrolyysituotteiden fermentointi bioetanoliksi ja biobutanoliksi. Olkeen perustuvalle bioetanolin ja biobutanolin tuotannolle tehtiin mallinnukseen perustuva kasvihuonekaasupäästöjen analyysi. Tanniinia ja bioetanolia tuottavalle kuoribiojalostamolle laadittiin aine- ja energiataseet kuvaava prosessimalli, jonka taloudellisuutta tutkittiin edelleen parametrisen herkkyysanalyysin avulla.
ISBN, ISSN	ISBN 978-951-38-7902-0 (URL: http://www.vtt.fi/publications/index.jsp) ISSN 2242-122X (URL: http://www.vtt.fi/publications/index.jsp)
Julkaisuaika	Joulukuu 2012
Kieli	English
Sivumäärä	159 s. + liitt. 17 s.
Projektin nimi	BioScen
Toimeksiantajat	
Avainsanat	Biorefinery, material properties, modelling, parameter optimisation, pyrolysis oil condensation, lignocellulosic biomass, hydrolysis, fermentation, greenhouse gas emissions
Julkaisija	VTT PL 1000, 02044 VTT, Puh. 020 722 111

BIOSCEN
Modelling Biorefinery Scenarios

The report includes major results from the research project BioScen – Modelling Biorefinery Scenarios (2008–2012), part of the Tekes funded BioRefine technology programme. BioScen was organised as a joint project of Aalto University (Aalto), University of Jyväskylä (JYU) and VTT. Industry support was from Neste Oil Oyj, Metso Power Oy, Pöyry Engineering Oy and Vapo Oy. The research goal was to develop material property and algorithmic basis for the simulation of the unit processes of the future biorefinery. The work covered physical, chemical and transport properties of typical biorefinery chemicals, parameter optimisation with sensitivity analysis, flowsheet modelling and analysis of the greenhouse gas emissions of selected biorefining products.

ISBN 978-951-38-7902-0 (URL: http://www.vtt.fi/publications/index.jsp)
ISSN 2242-122X (URL: http://www.vtt.fi/publications/index.jsp)