Dynamic modeling and validation of a lignocellulosic enzymatic hydrolysis process – A demonstration scale study

ARTICLE *in* BIORESOURCE TECHNOLOGY · OCTOBER 2013

Impact Factor: 4.49 · DOI: 10.1016/j.biortech.2013.10.029 · Source: PubMed

2 AUTHORS:

Remus Mihail Prunescu
Technical University of Denmark
7 PUBLICATIONS **17** CITATIONS
SEE PROFILE

Gürkan Sin
Technical University of Denmark
184 PUBLICATIONS **1,798** CITATIONS
SEE PROFILE

Bioresource Technology 150 (2013) 393–403

Contents lists available at ScienceDirect

Bioresource Technology

journal homepage: www.elsevier.com/locate/biortech

Dynamic modeling and validation of a lignocellulosic enzymatic hydrolysis process – A demonstration scale study

Remus Mihail Prunescu [a], Gürkan Sin [b,*]

[a] Department of Electrical Engineering, Automation and Control Group, Technical University of Denmark, Elektrovej Building 326, 2800 Kgs. Lyngby, Denmark
[b] CAPEC, Department of Chemical and Biochemical Engineering, Technical University of Denmark, Søltofts Plads Buildings 227 and 229, 2800 Kgs. Lyngby, Denmark

HIGHLIGHTS

- A dynamic model for demonstration scale liquefaction processes.
- The model includes a competitive cellulose and xylan hydrolysis mechanism.
- The model couples pH and viscosity calculations.
- Real biorefinery measurements are used to validate the model.
- Uncertainty and sensitivity analysis considering kinetic and feed composition.

ARTICLE INFO

Article history:
Received 17 July 2013
Received in revised form 2 October 2013
Accepted 9 October 2013
Available online 17 October 2013

Keywords:
Biofuel
Enzymatic hydrolysis
Demonstration scale
Dynamic modeling
Sensitivity and uncertainty analysis

ABSTRACT

The enzymatic hydrolysis process is one of the key steps in second generation biofuel production. After being thermally pretreated, the lignocellulosic material is liquefied by enzymes prior to fermentation. The scope of this paper is to evaluate a dynamic model of the hydrolysis process on a demonstration scale reactor. The following novel features are included: the application of the Convection–Diffusion–Reaction equation to a hydrolysis reactor to assess transport and mixing effects; the extension of a competitive kinetic model with enzymatic pH dependency and hemicellulose hydrolysis; a comprehensive pH model; and viscosity estimations during the course of reaction. The model is evaluated against real data extracted from a demonstration scale biorefinery throughout several days of operation. All measurements are within predictions uncertainty and, therefore, the model constitutes a valuable tool to support process optimization, performance monitoring, diagnosis and process control at full-scale studies.

1. Introduction

Bioethanol is thought to become a viable alternative to fossil fuels (Datta et al., 2011) and many countries with agricultural resources show an increasing interest in the second generation biofuel production technology, especially in USA, Brazil, Denmark and Italy. The latest developments that approach biorefineries to a commercial reality are presented by Larsen et al. (2012).

In biorefinery concepts that employ a biochemical conversion route, the lignocellulosic material originating from agricultural wastes is typically transformed into bioethanol following 4 major steps, i.e. pretreatment, enzymatic hydrolysis, fermentation and separation (Larsen et al., 2008). The pretreated lignocellulosic material has a high content of cellulose, hemicellulose and lignin. The purpose of the enzymatic hydrolysis process is to break down the long polymeric chains of cellulose and xylan into its constituent sugar units, mostly glucose and xylose.

The underlying mechanism of the enzymatic hydrolysis process is a complex one, which has been the target of many modeling studies. Kadam et al. (2004) formulates a mechanistic conversion model while Zhang and Lynd (2004) explain in detail how the cellulosic depolymerization occurs under the effect of enzymes. The Kadam model has been validated in laboratory scale experiments both by Kadam et al. (2004) and Hodge et al. (2009). Sin et al. (2010) have performed an identifiability and uncertainty analysis of the Kadam model and Morales-Rodriguez et al. (2011) have integrated it into a dynamic modeling framework for assessing biorefinery configurations. The depolymerization model has been extended and validated in laboratory scale experiments by Hosseini and Shah (2011a,b). So far, none of these models have been evaluated at a realistic demonstration scale.

In order to formulate a model for large scale biorefineries, several extensions need to be made to the previous models. pH calculations were not performed before although it is well known that

* Principal corresponding author. Tel.: +45 45252806.
E-mail addresses: rmpr@elektro.dtu.dk (R.M. Prunescu), gsi@kt.dtu.dk, gurkansin@gmail.com (G. Sin).

0960-8524/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.biortech.2013.10.029

Nomenclature

α_i	enzyme cocktail composition [–]
β	acetic acid to xylose ratio [–]
β_i	standardized regression coefficients given by the SRC method
β_{i_C}	SRC coefficients related to cellobiose output
β_{i_G}	SRC coefficients related to glucose output
β_{i_X}	SRC coefficients related to xylose output
δ_k^{msqr}	delta mean square sensitivity measure of parameter k [–]
ϵ_C	NIR error related to cellulose content [%]
ϵ_F	sensor error related to inflow of fiber mass [kg s^{-1}]
ϵ_X	NIR error related to xylan content [%]
η_i	temperature and pH dependency of r_i [0–1]
μ_l	liquid viscosity [kg m^{-1} s^{-1}]
μ_r	relative viscosity [–]
μ_s	slurry viscosity [kg m^{-1} s^{-1}]
μ_W	water viscosity [kg m^{-1} s^{-1}]
∇	gradient operator [–]
Φ	volume fraction of solids [–]
ρ	molecular density [g cm^{-3}]
σ_{θ_i}	parameter standard deviation used in the SRC method
σ_y	Monte-Carlo model output standard deviation used in the SRC method
A_G	viscosity fitting parameter for glucose [g m^{-1} s^{-1}]
A_T	concentration of acetate species inside the reactor [mol L^{-1}]
A_W	viscosity fitting parameter for pure liquid [g m^{-1} s^{-1}]
B_G	viscosity fitting parameter for glucose [K]
b_i	regression coefficients used in the SRC method
B_T	concentration of basic species inside the reactor [mol L^{-1}]
B_W	viscosity fitting parameter for pure liquid [K]
C	species concentration [g kg^{-1}] or [mol L^{-1}]
C_0	pretreated fiber chemical composition [g kg^{-1}]
C_B	base concentration [g kg^{-1}]
C_E	enzyme concentration [g kg^{-1}]
C_T	concentration of carbonic species inside the reactor [mol L^{-1}]
C_{C_S}	cellulose concentration [g kg^{-1}]
C_C	cellobiose concentration [g kg^{-1}]
C_{E_1}	CBH + EG concentration [g kg^{-1}]
C_{E_2}	G concentration [g kg^{-1}]
C_{E_3}	X concentration [g kg^{-1}]
C_{E_4}	concentration of other type of enzymes [g kg^{-1}]
$C_{E_{3_B}}^X$	concentration of hemicellulases bounded to xylan [g kg^{-1}]
$C_{E_{i_B}}$	bounded enzymes concentration [g kg^{-1}]
$C_{E_{i_F}}$	free enzymes concentration [g kg^{-1}]
$C_{E_{i_B}}^C$	concentration of type i enzymes bounded to cellulose [g kg^{-1}]
C_E	enzymes concentration [g kg^{-1}]
C_F	furfural concentration [g kg^{-1}]
C_G	glucose concentration [g kg^{-1}]
C_{K_0}	inflow concentration of reactor species [g kg^{-1}]
C_K	array with all reactor species [g kg^{-1}]
C_{L_S}	lignin concentration [g kg^{-1}]
C_S	total concentration of solids [g kg^{-1}]
C_{X_S}	xylan concentration [g kg^{-1}]
C_X	xylose concentration [g kg^{-1}]
D	diffusion coefficient [m^2 s^{-1}]
E_{M_i}	maximum adsorbed enzymes [–]
F_μ	uncertainty feed parameter related to mean value distribution
F_σ	uncertainty feed parameter related to standard deviation distribution
F_B	base inflow [kg s^{-1}]
F_E	enzyme inflow [kg s^{-1}]
F_{FF}	pretreated fiber inflow [kg s^{-1}]
F_{FM}	fiber mash outflow [kg s^{-1}]
F_W	water inflow [kg s^{-1}]
I_3	overall inhibition term for r_3 [g kg^{-1}]
I_{C_i}	inhibition of r_i by cellobiose [g kg^{-1}]
I_{F_i}	inhibition of r_i by furfural [g kg^{-1}]
I_{G_i}	inhibition of r_i by glucose [g kg^{-1}]
I_{X_i}	inhibition of r_i by xylose [g kg^{-1}]
k	Boltzmann constant [J K^{-1}]
K_i	reaction rate constant of r_i [kg g^{-1} s^{-1}]
K_{A_i}	adsorption constant [–]
$k_{G\mu}$	glucose correlation parameter [m^2 s^{-1}]
$K_{La_{CO_2}}$	CO_2 mass transfer coefficient [d^{-1}]
L_R	reactor length [m]
L_T	concentration of lactic species inside the reactor [mol L^{-1}]
M	molecular mass [g mol^{-1}]
M_{FM}	reactor holdup [kg]
N	Avogadro number [mol^{-1}]
pH	fiber mash pH [–]
R	production or consumption [g kg^{-1} s^{-1}] or [mol L^{-1} s^{-1}]
r	molecular radius [Å]
r_1	cellulose to cellobiose rate [g kg^{-1} s^{-1}]
r_2	cellulose to glucose rate [g kg^{-1} s^{-1}]
r_3	cellobiose to glucose rate [g kg^{-1} s^{-1}]
r_4	xylan to xylose rate [g kg^{-1} s^{-1}]
r_5	Cellobiose to glucose rate [g kg^{-1} s^{-1}]
r_{CTR}	CO_2 stripping rate [mol L^{-1} s^{-1}]
S_T	concentration of succinic species inside the reactor [mol L^{-1}]
$s_{nd,ik}$	non dimensional sensitivity of output i with respect to parameter k [–]
T	absolute temperature [K]
t_R	reactor retention time [s]
V	molecular volume [cm^3]
v	velocity vector [m s^{-1}]
v_x	horizontal speed [m s^{-1}]
y_{reg}	linear model fit used in the SRC method
CO_2^*	atmospheric CO_2 concentration [mol L^{-1}]

enzymes are sensitive to pH following a Gaussian curve. At laboratory scale, perfect pH control can be easily achieved and assumed but at larger scales such control becomes challenging. When the process runs in a continuous manner, the inflow composition varies due to different pretreatment conditions or biomass composition and the concentration of acetic acid has disturbances that affect the pH level (Prunescu et al., 2013). Also, previous models do not include xylan hydrolysis, which contributes to acetic acid formation inside the reactor leading to a pH profile along the container that affects the enzymatic activity. Another missing feature is viscosity calculations, which are important for estimating diffusion coefficients and for assessing transportation and mixing effects.

With these in background, a dynamic lignocellulosic hydrolysis model adequate for demonstration and large scale processes is formulated with the following features: the reactor is modeled as plug flow using the Convection–Diffusion–Reaction (CDR) equation in one dimension in order to capture properly transport and

mixing effects along the length of the tank, i.e. the x axis; the reaction kinetics comprises competitive cellulose and xylan hydrolysis with temperature and pH dependency plus furfural inhibition; pH dynamics account for most of the relevant buffers that come from the pretreatment process, i.e. acetic, succinic and lactic acids plus a bicarbonate formation system with CO_2 stripping effects. Chemical kinetics, and pH and viscosity calculations are tracked in each cell of the 1-D CDR model.

The model is then analyzed in order to assess its sensitivity to kinetic and feed parameters, and to quantify its predictions uncertainty.

This paper is structured as follows: Section 2 describes a typical demonstration scale biorefinery and how measurements were recorded for comparing against model predictions; Section 3 formulates the mathematical model and includes its sensitivity and uncertainty analysis; and Section 4 illustrates and discusses the model validation results.

2. Methods

A demonstration scale biorefinery is capable of transforming approximately 4 th^{-1} of biomass into three products, i.e. 576 kg h^{-1} 2G bioethanol, 1484 kg h^{-1} C5 molasses and 1740 kg h^{-1} lignin pallets (Larsen et al., 2012). The biomass can be any lignocellulosic material, preferably agricultural wastes, e.g. straw, bagasses or corn stover, or forest wastes, e.g. saw dust, pulp waste or thinned wood (Naik et al., 2010). There are various biomass pretreatment methods, e.g. autohydrolysis, steam explosion, acid hydrolysis etc., among which steam explosion is seen as a cheap and environmentally friendly pretreatment (Chiaramonti et al., 2012). In this investigation, wheat straw was pretreated with steam in a large pressurized thermal reactor described in Prunescu et al. (2012).

Fig. 1 illustrates the biomass conversion route with emphasis on the enzymatic hydrolysis process. The inflow of pretreated fibers is denoted as F_{FF} and has an initial chemical composition C_0, which is measured with online NIR equipment that performs a measurement every 5 min. A strong base, i.e. NaOH, is pumped into the tank near the inlet for pH adjustments. The base inflow F_B and its concentration C_B are considered known. A certain amount of enzymes proportional to cellulose content from the pretreated fibers is added by a pump from a storage tank. The flow of enzymes F_E and its concentration C_E are considered known. Water is also added with a known flow rate F_W. The reactor is mass controlled using a mass measurement M_{FM} and by manipulating the outflow of fiber mash F_{FM}. The outflow concentration C is determined by grabbing a sample from the outflow every 6 h and analyzing it in the laboratory with HPLC equipment. The fiber mash pH is measured online with a sampling period of 10 s and all inflows are measured with a sampling period of 2 s.

3. Model development

For this investigation, the mechanistic approach suggested by Kadam et al., 2004 is preferred as a starting base. This semi-mechanistic model seems suitable for the purpose of process operation studies as opposed to the more complex and detailed depolymerization type models (Sin et al., 2010).

3.1. Total mass balance

All flows are measured and a total mass balance can be easily constructed:

$$\frac{dM_{FM}}{dt} = F_W + F_{FF} + F_E + F_B - F_{FM} \tag{1}$$

3.2. Transport modeling

The pretreatment process creates a slurry with 25–30% of solid particles, i.e. cellulose, xylan and lignin (Jørgensen et al., 2007). The viscosity of this slurry is relatively high in the first phase of the hydrolysis leading to a plug flow transport phenomena in the liquefaction reactor. The container is assumed to have a much larger

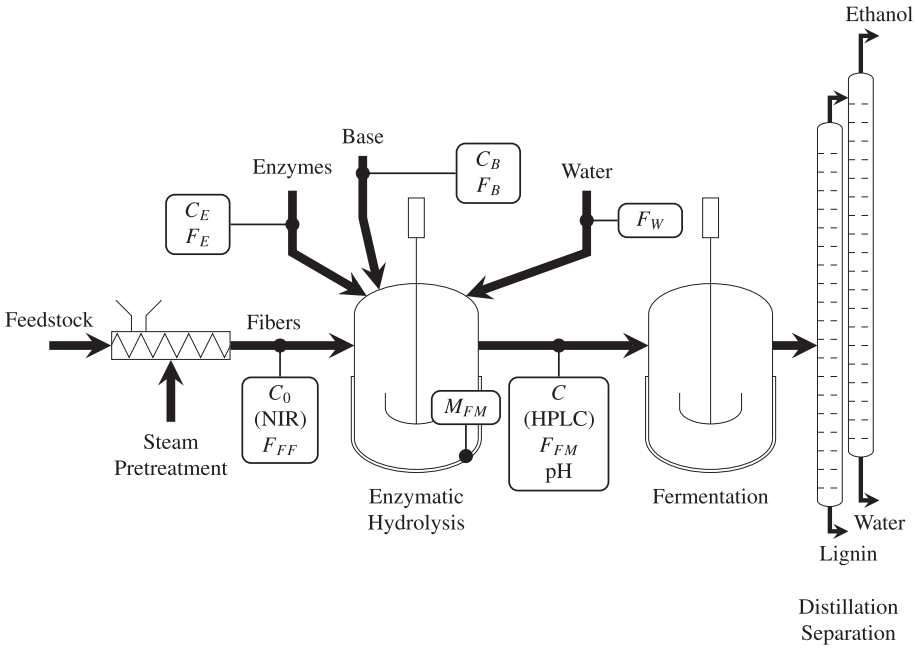

Fig. 1. Biorefinery setup with focus on the enzymatic hydrolysis reactor. All flows are measured, i.e. pretreated fibers F_{FF}, base F_B, water F_W, enzymes F_E and fiber mash F_{FM}. Inflow concentration C_0 is measured by online NIR equipment while C_B and C_E are considered known. There are 2 additional measurements, i.e. a pH sensor on the outflow and a mass indicator M_{FM} for the reactor holdup. The outstream concentration C is determined by analyzing fiber mash samples in the laboratory with an HPLC device.

length than its height, the slurry traversing the tank horizontally from left to right. The mixture is assumed to have uniform properties along the height of the reactor due to vertical mixers.

The species concentration change with respect to time is a combination of convection and diffusion effects plus production or consumption terms. The process normally runs in a continuous manner and slurry transportation is dominated by convection effects. However, sometimes it is necessary to stop the outflow of the tank due to mechanical faults, for example. In such a case, the hydrolysis switches to batch mode and diffusion effects become prominent as cellulosic conversion continues. Convection and diffusion effects are well captured by the Convection–Diffusion–Reaction mass conservation equation (Bird et al., 2006):

$$\frac{\partial C}{\partial t} = -\nabla(\vec{v}C) + \nabla(D\nabla C) + R \tag{2}$$

where C is the species concentration, $\vec{v}$ is the speed vector, D is the diffusion coefficient and R is the production or consumption rate presented in detail in subsection 3.4. The term containing $\vec{v}$ indicates the concentration change due to advection while the term containing D indicates the concentration rate due to diffusion effects. The advection term appears negative because the axis orientation is positively aligned with $\vec{v}$ in the continuity equation. The gradient operator ∇ from Eq. (2) contains the derivatives only along the x axis due to the vertical uniformity assumption. In this case, the diffusion part of Eq. (2) is found according to Fick's second law (Bird et al., 2006):

$$\nabla(D\nabla C) = \frac{\partial}{\partial x}\left(D\frac{\partial C}{\partial x}\right) \tag{3}$$

Substituting Eq. (3) in (2) yields:

$$\frac{\partial C}{\partial t} = -v_x\frac{\partial C}{\partial x} + \frac{\partial}{\partial x}\left(D\frac{\partial C}{\partial x}\right) + R \tag{4}$$

Eq. (4) is solved following a finite element method specific to computational fluid dynamics. Briefly, the reactor is divided into n cells and the space partial derivatives from Eq. (4) are approximated using a finite volume method as described in (Egeland and Gravdahl, 2002) and given in supplementary material.

The horizontal speed v_x is considered constant along the length of the reactor and is computed as follows:

$$v_x = \frac{L_R}{t_R} \tag{5}$$

where L_R is the length of the liquefaction reactor and t_R is the retention time. The retention time is the ratio between total mass and outflow:

$$t_R = \frac{M_{FM}}{F_{FM}} \tag{6}$$

where M_{FM} is the total mass of fiber mash inside the tank and F_{FM} is the outflow.

Diffusion occurs only in the liquid fraction of the slurry and affects only soluble particles, not solids. It is expected to have a varying diffusion along the x axis because the slurry viscosity is changing (expected to decrease) as the liquefaction progresses while liquid viscosity is expected to increase as sugars are formed and dissolved. The Stokes–Einstein equation describes the diffusion coefficient D of solubles as a function of temperature T and liquid viscosity μ_l:

$$D = \frac{kT}{6\pi r\mu_l} \tag{7}$$

where k is the Boltzmann constant and r is the molecular radius of the dissolved particle. The radius r of the molecule can be approximated with the following relation:

$$M = N\rho V = N\rho\left(\frac{4}{3}\pi r^3\right) \Rightarrow r = \left(\frac{3M}{4\pi N\rho}\right)^{\frac{1}{3}} \tag{8}$$

where M is the molecular mass of soluble component, N is the Avogadro's number, ρ is the solute density and V is its molecular volume. The liquid viscosity μ_l remains to be computed in order to solve Eq. (7). An empirical relative viscosity equation for high concentration slurries was derived by Thomas (1965):

$$\mu_r = a_0 + a_1\Phi + a_2\Phi^2 \tag{9}$$

where Φ is the volume fraction of solid particles and a_1 and a_2 are calibration coefficients estimated as follows by Thomas (1965):

$$a_0 = 1 \qquad a_1 = 2.5 \qquad a_2 = 10.05 \tag{10}$$

The relative viscosity μ_r and the liquid viscosity μ_l can then be used to compute the slurry viscosity μ_s:

$$\mu_s = \mu_r\mu_l \tag{11}$$

As glucose is produced and dissolved in the liquid part of the slurry, the liquid viscosity changes following a simple linear relation as given by Converti et al. (1999):

$$\mu_l = \mu_W + C_G k_{G\mu} \tag{12}$$

where μ_W is the viscosity of the pure liquid solvent, C_G is the glucose concentration and $k_{G\mu}$ is a correlation parameter for glucose. $k_{G\mu}$ and μ_W follow an empirical exponential law also known as the Guzman–Andrade equation (Glasstone et al., 1941):

$$\mu_W = A_W e^{\frac{B_W}{T}} \qquad k_{G\mu} = A_G e^{\frac{B_G}{T}} \tag{13}$$

where A_W and B_W are fitting parameters for pure liquid, T is the reactor temperature, and A_G and B_G are parameters for sugar correlation. μ_W is approximately water viscosity and numerical values for A_W, B_W, A_G and B_G are given by Converti et al. (1999):

$$\begin{aligned} A_W &= 2.41\times10^{-3}\mathrm{gm^{-1}s^{-1}} \\ B_W &= 1774.9\mathrm{K} \\ A_G &= 8.65\times10^{-10}\mathrm{m^2s^{-1}} \\ B_G &= 2502\mathrm{K} \end{aligned} \tag{14}$$

3.3. pH modeling

The pH model is implemented as a set of algebraic equations coupled with the CDR model (Eq. 4). The model is an extension of the study presented in an earlier publication (Prunescu et al., 2013) and comprises a set of 8 weak acids and strong base hydration reactions: self-ionization of water (Eq. (15c)), acetic acid dissociation (Eq. (15a)), carbonic acid formation and dissociation (Eqs. (15d)), succinic acid dissociation (Eqs. (15f) and (15g)) and lactic acid dissociation (Eq. (15h)). The hydration equations are modeled using the approach described by McAvoy et al. (1972). The complete set of equilibrium dissociation equations is shown next:

$$C_2H_4O_2 \rightleftharpoons C_2H_3O_2^- + H^+ \tag{15a}$$
$$NaOH \rightarrow Na^+ + OH^- \tag{15b}$$
$$H^+ + OH^- \rightleftharpoons H_2O \tag{15c}$$
$$CO_2 + H_2O \rightleftharpoons HCO_3^- + H^+ \tag{15d}$$
$$HCO_3^- \rightleftharpoons CO_3^{2-} + H^+ \tag{15e}$$
$$C_4H_6O_4 \rightleftharpoons C_4H_5O_4^- + H^+ \tag{15f}$$
$$C_4H_5O_4^- \rightleftharpoons C_4H_4O_4^{2-} + H^+ \tag{15g}$$
$$C_3H_6O_3 \rightleftharpoons C_3H_5O_3^- + H^+ \tag{15h}$$

Due to the fact that the reactor tank is exposed to atmospheric pressure, it is expected to have a CO_2 stripping process from liquid to gas phase that leads to changes in the carbonic acid and bicarbonate buffers. The CO_2 stripping process or *CTR* is modeled as in Sin and Vanrolleghem (2007) with typical saturation kinetics:

$$CTR = K_{La_{CO_2}}\left(CO_2^* - CO_{2_{[aq]}}\right) \tag{16}$$

where $K_{La_{CO_2}}$ is the mass transfer coefficient, CO_2^* is the atmospheric CO_2 concentration and $CO_{2_{[aq]}}$ is the dissolved CO_2 concentration in the reactor.

The dissociation constants for each species at 50 °C, which is a typical optimal reactor temperature for an enzymatic process, are given in Table 1.

The pH system of equations is solved in each cell of the grid following the procedure from (Prunescu et al., 2013). Thus, the concentration of H^+ is found, which is necessary to compute the pH level:

$$pH = -\log_{10}[H^+] \tag{17}$$

3.4. Modeling of cellulose and xylan conversion

The kinetic model is an extension of the work performed by Kadam et al. (2004) which is used here to calculate the reaction term *R* in the CDR model (Eq. 4). Here, the Kadam model is extended taking into consideration enzymes that are capable of decomposing both celullose and xylan, inhibition by furfural of the enzymatic activity, and pH dependency of the enzymatic activity. During xylan hydrolysis, acetic acid is released contributing to lowering the pH level in the mixture. Tracking the acetic acid formation is crucial for building the pH profile of the reactor.

The overall conceptual hydrolysis mechanism is shown in Fig. 2. Cellulose is decomposed into cellobiose and glucose with reaction rates r_1 and r_2, respectively. Cellobiose is further decomposed into glucose with rate r_3. Xylan is hydrolyzed into xylose with an r_4 rate and acetic acid is released as a by-product with an r_5 rate. Cellobiose and glucose formation are inhibited by the sugars themselves or by furfural, which was formed in the pretreatment process.

The following chemical species are tracked in the conversion model:

$$C_K = [\,C_{C_S} \quad C_{X_S} \quad C_{L_S} \quad C_C \quad C_G \quad C_X \quad C_F \quad C_E\,]^\top \tag{18}$$

where C_{C_S}, C_{X_S} and C_{L_S} are solid components from the slurry, i.e. cellulose, xylan and lignin concentrations, and C_C, C_G, C_X, C_F and C_E constitute the liquid part of the slurry, i.e. cellobiose, glucose, xylose, furfural and enzymes, respectively. Eq. (4) is evaluated for each species from Eq. (18). The boundary conditions of Eq. (4) are constructed from the initial chemical composition of pretreated fibers, or C_{K_0}, which is measured in reality with NIR equipment.

The enzymes cocktail is a complex mixture containing many protein types that could be divided into cellulase, hemicellulase and other type of proteins. Cellulose is hydrolyzed by the cellulase group, which is subdivided into exo-β-1.4-cellobiohydrolase and endo-β-1.4-glucanase (CBH + EG), notated as C_{E_1}, and β-glucosidase (G) enzymes, i.e. C_{E_2}. Xylan is hydrolyzed by the hemicellulase (X) group, i.e. C_{E_3}. There are other types of enzymes in the mixture in negligible amounts and their concentration is notated as C_{E_4}. The enzymatic cocktail can then be parametrized in α_i with i from 1 to 4 representing the fraction of each enzyme type. The total enzyme concentration can then be expressed as the sum of each enzyme concentration:

$$C_E = C_{E_1} + C_{E_2} + C_{E_3} + C_{E_4} \tag{19}$$

Table 1
Chemical compounds relevant to pH calculation with their dissociation or hydration constants at 50 °C.

Formula	Description	Value	Reference	Variable
$C_2H_4O_2$	Acetic acid	1.63e-5	Albert and Serjeant (1962)	K_A
H_2O	Water	5.39e-14	Albert and Serjeant (1962)	K_W
H_2CO_3	Carbonic acid	5.14e-7	Albert and Serjeant (1962)	K_{C_1}
HCO_3^-	Bicarbonate	6.69e-11	Albert and Serjeant (1962)	K_{C_2}
$C_4H_6O_4$	Succinic acid	6.51e-5	Albert and Serjeant (1962)	K_{S_1}
$C_4H_5O_4^-$	Succinate ion	2.08e-6	Albert and Serjeant (1962)	K_{S_2}
$C_3H_6O_3$	Lactic acid	1.27e-4	Albert and Serjeant (1962)	K_L

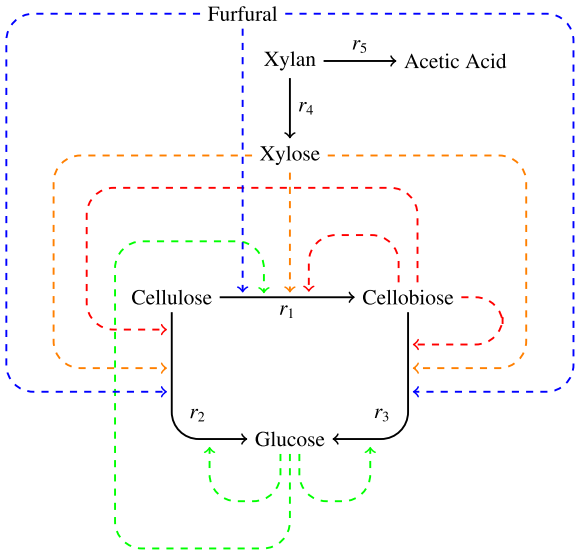

Fig. 2. Conceptual hydrolysis mechanism extended from Kadam et al., 2004. Cellulose is converted by CBH + EG and G into cellobiose and glucose with reaction rates r_1 and r_2, respectively. Cellobiose is converted into glucose with rate r_3. Hemicellulose hydrolysis is illustrated by xylan conversion into xylose with rate r_4, which also releases acetic acid with rate r_5. Dashed lines show inhibition by sugars and furfural (extended from Kadam et al., 2004).

There is an equilibrium between bounded and free enzymes, which is simplified to the following equation:

$$\text{Enzyme}_F + \text{Solids} \leftrightarrow \text{Enzyme-Solids}_B \tag{20}$$

where Enzyme_F are free enzymes and Enzyme-Solids_B are bounded enzymes to solids. Therefore, enzymes can be in one of the two states: bounded to solids (B) or free (F). The total concentration of CBH + EG, G or X enzymes consists of bounded and free enzymes:

$$C_{E_i} = C_{E_{i_B}} + C_{E_{i_F}} \tag{21}$$

where $i = 1, 2, 3$. The effects of the other type of enzymes, i.e. $i = 4$, is neglected as they do not participate actively to cellulose and xylan decomposition.

The equilibrium between adsorbed and free enzymes follows a Langmuir isotherm expression. In the case of type i enzymes, where i can be CBH + EG, G or X, the Langmuir isotherm relation expresses the ratio between concentration of adsorbed enzymes and solids:

$$\frac{C_{E_{i_B}}}{C_S} = E_{M_i}\frac{K_{A_i}C_{E_{i_F}}}{1 + K_{A_i}C_{E_{i_F}}} \tag{22}$$

where $C_{E_{i_B}}$ represents the bounded enzymes, C_S is the total concentration of solids, E_{M_i} is the maximum adsorbed enzymes, K_{A_i} is the adsorption constant and $C_{E_{i_F}}$ is the free enzymes concentration. Cellulose is a percentage of the total solids from the mixture. The concentration of adsorbed enzymes to cellulose $C^C_{E_{i_B}}$ can then be calculated as:

$$C^{C}_{E_{i_B}} = C_{E_{i_B}} \frac{C_{C_S}}{C_S} = E_{M_i} \frac{K_{A_i} C_{E_{i_F}}}{1 + K_{A_i} C_{E_{i_F}}} C_{C_S} \tag{23}$$

where C_{C_S} is the substrate or cellulose concentration for $i = 1, 2$. Similar computations are performed in the case of hemicellulases, or $i = 3$, but the substrate is the xylan concentration this time, or C_{X_S}:

$$C^{X}_{E_{3_B}} = E_{M_3} \frac{K_{A_3} C_{E_{3_F}}}{1 + K_{A_3} C_{E_{3_F}}} C_{X_S} \tag{24}$$

where $C^{X}_{E_{3_B}}$ represents the concentration of enzymes bound to xylan.

The cellulose to cellobiose reaction rate r_1 with competitive glucose, cellobiose, xylose and furfural inhibition is developed starting from the expression given in Kadam et al., 2004 with the extension of pH dependency and furfural inhibition:

$$r_1 = \frac{K_1 \eta_1(T, pH) C^{C}_{E_{1_B}} C_{C_S}}{1 + \frac{C_C}{I_{C_1}} + \frac{C_X}{I_{X_1}} + \frac{C_G}{I_{G_1}} + \frac{C_F}{I_{F_1}}} \tag{25}$$

where K_1 is the reaction constant, $\eta_1(T, pH)$ is the temperature and pH dependency term, I_{C_1} is the cellobiose inhibition term, I_{X_1} is the xylose inhibition term, I_{G_1} is the glucose inhibition term, and I_{F_1} is the furfural inhibition term. Only CBH + EG participate in this reaction.

The cellulose to glucose reaction rate r_2 with competitive glucose, cellobiose, xylose and furfural inhibition is extended similarly:

$$r_2 = \frac{K_2 \eta_2(T, pH) \left(C^{C}_{E_{1_B}} + C^{C}_{E_{2_B}} \right) C_{C_S}}{1 + \frac{C_C}{I_{C_2}} + \frac{C_X}{I_{X_2}} + \frac{C_G}{I_{G_2}} + \frac{C_F}{I_{F_2}}} \tag{26}$$

where K_2 is the reaction constant, $\eta_2(T, pH)$ is the temperature and pH dependency term, I_{C_2} is the cellobiose inhibition term, I_{X_2} is the xylose inhibition term, I_{G_2} is the glucose inhibition term, and I_{F_2} is the furfural inhibition term. Both CBH + EG and G enzymes participate in this reaction.

The cellobiose to glucose reaction rate r_3 with competitive glucose, cellobiose, xylose and furfural inhibition is extended similarly starting form (Kadam et al., 2004):

$$r_3 = \frac{K_3 \eta_3(T, pH) C_{E_{2_F}} C_C}{I_3 \left(1 + \frac{C_X}{I_{X_3}} + \frac{C_G}{I_{G_3}} + \frac{C_F}{I_{F_3}} \right) + C_C} \tag{27}$$

where K_3 is the reaction constant, I_3 is an overall inhibition term and I_{X_3}, I_{G_3} and I_{F_3} represent inhibition from xylose, glucose and furfural, respectively.

Xylan decomposition is modeled similarly to reaction rate r_1 but with xylan substrate:

$$r_4 = \frac{K_4 \eta_4(T, pH) C^{X}_{E_{3_B}} C_{X_S}}{1 + \frac{C_C}{I_{C_4}} + \frac{C_X}{I_{X_4}} + \frac{C_G}{I_{G_4}} + \frac{C_F}{I_{F_4}}} \tag{28}$$

where r_4 is the xylan to xylose reaction rate, K_4 is the reaction constant, $\eta_4(T, pH)$ represents the temperature and pH dependency, $C^{X}_{E_{3_B}}$ is the concentration of bounded hemicellulases to xylan, and C_{X_S} is the xylan substrate. $I_{C_4}, I_{X_4}, I_{G_4}$ and I_{F_4} are inhibition terms for cellobiose, xylose, glucose and furfural, respectively.

The temperature and pH dependency can be usually retrieved from the enzymes supplier (e.g. Novozymes). If such data are available the following relation is proposed:

$$\eta_i(T, pH) = \eta_T(T) \eta_{pH}(pH) \tag{29}$$

where η_T and η_{pH} are efficiency factors with values between 0 to 1 related to temperature and pH, respectively. The efficiency factors can be either table based or following any arbitrary Gaussian curve functions. Two typical curves for pH and temperature dependency can be found in the Supplementary material. These dependencies are implemented as numerical tables and linear interpolation between data points is used throughout simulation.

The amount of acetic acid contained in the hemicellulose part depends on biomass type. For example, in the case of hardwood biomass, 0.5mol of acetic acid is released for every 1mol of xylose (Khan, 2010). Using the molaric mass of xylose and acetic acid, it is found that for every 1g of xylose, 0.2g of acetic acid is released. This ratio is parametrized with β and the acetic acid production rate r_5 becomes:

$$r_5 = \beta r_4 \tag{30}$$

The predictions of cellobiose, glucose and xylose concentrations are regarded as model outputs:

$$y = [C_C \quad C_G \quad C_X] \tag{31}$$

The complete model has 31 parameters: 4 that describe the enzyme mixture composition, i.e. α_i, 1 for the acetic acid to xylose ratio, i.e. β, and 26 parameters relevant for the enzymatic hydrolysis kinetics. α_i and β are fixed a priori and put into vector θ_F. The enzymatic kinetics parameters are gathered into vector θ_K.

3.5. Model calibration, sensitivity and uncertainty analysis

The previous analysis of the Kadam model performed by Sin et al., 2010 indicated that the model was over parameterized with respect to available measurements. Hence, for a proper model calibration, a sensitivity analysis was performed to find out the most significant parameters to be used as a subset for fine tuning of the model fits to the available data. The sensitivity analysis is conducted by calculating a measure called δ_k^{msqr} as in the methodology described in (Brun et al., 2001):

$$\delta_k^{msqr} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (s_{nd,ik})} \tag{32}$$

where k is the parameter index in the hydrolysis parameter vector θ_K, i is the model output index and $s_{nd,ik}$ is the non dimensional sensitivity defined as:

$$s_{nd,ik} = \frac{\partial y_i}{\partial \theta_k} \frac{\theta_k}{sc_i} \tag{33}$$

where $\partial y_i / \partial \theta_k$ represents the output variation with respect to a variation in parameter θ_k and sc_i is a scaling factor with the same physical dimension as the corresponding observation. All parameters are ranked according to δ_k^{msqr} and a subset θ_S is built with the most significant parameters.

A model calibration follows, which adjusts the parameters from the θ_S subset in order to obtain a better fit. For this calibration, only 50 h out of the 170 h of recorded data are used, while the remaining 120 h data are used for validation. It is noted that the calibration of the model parameters is meant to be a fine-tuning around the nominal values of the model parameters estimated from batch assays by Kadam et al. (2004). A comprehensive system identification is not pursued due to the fact that available plant measurements were obtained from closed loop operation and not under proper optimal experimental design.

3.5.1. Uncertainty analysis

The uncertainty analysis is carried out using the engineering standard Monte Carlo technique, which includes the following four steps (Morales-Rodriguez et al., 2012): (1) define input uncertainties with their range; (2) sampling of kinetics and feed parameters using the Latin hypercube sampling with correlation control;

(3) run Monte Carlo simulations with sampled values; (4) evaluate results. In the last step, in addition to inference statistics, a sensitivity analysis was also performed using linear regression of Monte Carlo outputs, also known as the standardized regression coefficients (SRC). The methodology is detailed in (Morales-Rodriguez et al., 2012).

In step 1, the uncertainty analysis considers two sources of uncertainty, namely kinetic parameters and feed composition (mass). For definition of the kinetic parameter uncertainty, a uniform distribution with the corresponding lower and upper bounds are used as defined in (Morales-Rodriguez et al., 2012), while the correlation matrix between the parameters was taken from (Sin et al., 2010). Bias and standard deviation of measurements due to sensors were considered for the characterization of feed composition uncertainty. The feed measurements are subject to errors either due to a miss calibration of the NIR equipment or because of a wrongly placed sensor, which often indicates offsets in the mass inflow of fibers. The most relevant feed variables are cellulose and xylan content, and fiber mass inflow. The error that affects these signals are statistically characterized using normal distributions with the following parameters:

$$\epsilon_C \in N(\mu_C, \sigma_C^2) \quad \epsilon_X \in N(\mu_X, \sigma_X^2) \quad \epsilon_F \in N(\mu_F, \sigma_F^2) \tag{34}$$

where ϵ_C, ϵ_X and ϵ_F are cellulose, xylan and inflow errors, and N is the normal distribution. The mean values μ_C, μ_X and μ_F represent sensor offsets and are normally distributed such that to cover a range of 5–10% of their nominal operational values, an error range assumed for NIR equipments. The standard deviations of the measurements are considered to follow gamma distributions, which is good practice in measurement error modeling. Hence the uncertainty on the measured feed composition are generated by performing LHS on 2 parameters of the normal distribution: one that characterizes the probability of the mean value, notated as F_μ, and another one for the standard deviation, i.e. F_σ.

The SRC method fits a linear regression model to Monte Carlo simulation outputs (Helton and Davis, 2003):

$$y_{reg} = b_0 + \sum_{i=1}^{N_\theta} b_i \theta_i \tag{35}$$

where θ_i is the uncertain parameter vector of length N_θ, b_i are the regression coefficients and y_{reg} is the fitted model output. The regression coefficients b_i are then scaled with respect to the standard deviation in parameters and simulation outputs:

$$\beta_i = \frac{\sigma_{\theta_i}}{\sigma_y} b_i \tag{36}$$

where β_i are the standardized regression coefficients, σ_{θ_i} is the standard deviation of parameter uncertainty and σ_y is the simulation output standard deviation.

3.6. Simulation scenarios

The model is evaluated in two different scenarios: the first one concerns a steady-state average performance simulation while in the second one the model is driven by real measurements and its predictions are compared against HPLC data throughout 170 h of operation.

In the first scenario, the reactor is assumed to be filled initially with water and all model inputs, including feed composition, are set to a typical operational point of a demonstration scale biorefinery as in Table 2. The reactor is capable of processing 4 th^{-1} of fibers. Water and enzymes inflows are adjusted proportionally to fibers inflow. A mass controller automatically sets the outflow of fiber mash in order to keep a constant holdup. The purpose of this scenario is to observe the steady-state concentrations, pH, viscosity and solids profiles from the reactor.

Table 2
Dry matter composition of steam pretreated wheat straw at 170 °C: demonstration scale plant values against composition determined by Ballesteros et al. (2006).

Component	Demonstration plant (%)	Ballesteros et al. (2006) (%)
Cellulose	45.0	50.0
Xylan	10.0	13.6
Lignin	32.0	20.8
Other	13.0	15.6

4. Results and discussion

4.1. Model sensitivity and calibration

Model parameters are initialized to reported values from (Kadam et al., 2004), which are given in Table 3. There are a few missing parameters referring to xylan hydrolysis and furfural inhibition, which have not been reported earlier in the literature. These parameters are initialized to reasonable values.

The δ^{msqr} sensitivity measure is numerically computed for each output as $\delta_C^{msqr}, \delta_G^{msqr}$ and δ_X^{msqr}, representing cellobiose, glucose and xylose, respectively, and also as an overall indicator δ^{msqr}. Parameters are ranked with respect to δ^{msqr} and the results are shown in Fig. 3. As expected, the model is over parametrized and only a subset of parameters is relevant for the output dynamics. A threshold for δ^{msqr} is set to 0.025 in this case (there is no absolute threshold value reported (Sin et al., 2010)) and shown in the bottom plot of Fig. 3. The subset of significant parameters is found as follows:

$$\theta_S = \begin{bmatrix} K_1 & K_2 & K_4 & E_{M_1} & K_{A_1} & I_{C_1} & I_{G_1} & I_{X_1} & I_{G_2} & I_{X_2} & I_{G_4} \end{bmatrix}^\top \tag{37}$$

Table 3
Model parameters initialization and calibration.

Parameter	Value	Kadam et al. (2004)
θ_F	Independently fixed parameters	
α_1	0.5	–
α_2	0.3	–
α_3	0.2	–
α_4	0	–
β	0.2	–
θ_S	Calibrated parameters	
K_2	0.0053	0.002
E_{M_1}	0.015	0.06
I_{X_2}	0.029	0.2
K_1	0.00034	0.0062
K_4	0.0027	–
I_{C_1}	0.0014	0.015
I_{G_4}	2.39	–
I_{G_1}	0.073	0.1
K_{A_1}	0.84	0.4
I_{G_2}	0.34	0.04
I_{X_1}	0.1007	0.1
Parameters with low sensitivity		
E_{M_2}	0.01	0.01
E_{M_3}	0.01	–
K_{A_2}	0.1	0.1
K_{A_3}	0.1	–
I_{F_1}	10	–
I_{C_2}	132	132
I_{F_2}	10	–
I_3	24.3	24.3
I_{X_3}	201	201
I_{G_3}	3.9	3.9
I_{F_3}	10	–
I_{C_4}	24.3	–
I_{X_4}	201	–
I_{F_4}	10	–
K_3	0.07	0.07

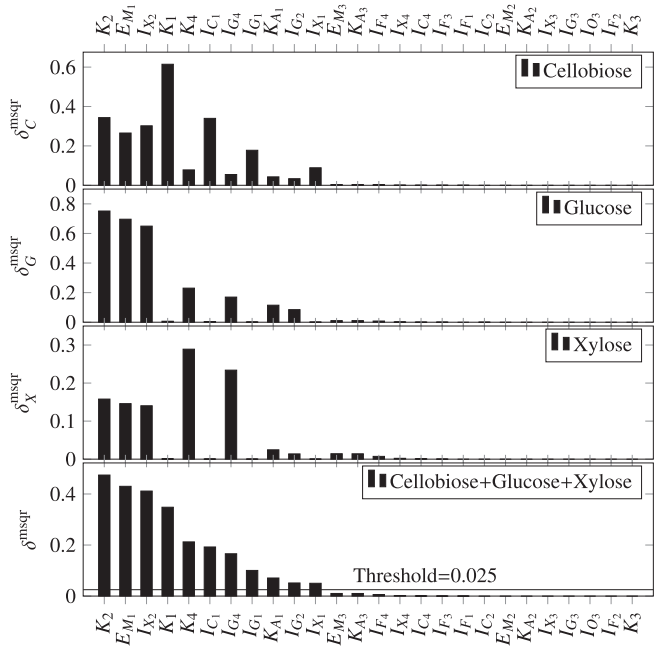

Fig. 3. Sensitivity measure δ^{msqr} of model outputs with respect to kinetics parameters θ_K. The first 3 plots show δ^{msqr} for cellobiose (δ_C^{msqr}), glucose (δ_G^{msqr}) and xylose (δ_X^{msqr}). The bottom plot illustrates the overall sensitivity measure δ^{msqr} and a threshold that delimits the relevant parameters subset θ_S from the rest of the parameters.

The sensitivity analysis results indicate that the reaction rates that participate in cellobiose, glucose and xylose formation, i.e. K_1, K_2 and K_4, are influential parameters. Inhibition of r_1 and r_2 by glucose and xylose, i.e. $I_{G_1}, I_{G_2}, I_{X_1}$ and I_{X_2}, as well as I_{G_4} or inhibition of r_4 by glucose, and I_{C_1} or inhibition of r_1 by cellobiose, also constitute important model parameters. The cellullase adsorption parameters E_{M_1} and K_{A_1} are the parameters closest to the threshold indicating a slight significance. The furfural inhibition parameters, i.e. $I_{F_1}, I_{F_2}, I_{F_3}$ and I_{F_4} have low sensitivity on model outputs, which is in accordance to (Cantarella et al., 2004) who observed a reduced inhibitory effect of furfural in the enzymatic hydrolysis process. All other model parameters, i.e. the ones that influence the cellobiose to glucose reaction r_3, and cellobiose inhibition on the other reactions, have almost no sensitivity on model outputs.

The parameters in θ_S are calibrated on a reduced set of recorded data in order to obtain a better fit. The calibrated values are shown in the second column of Table 3. As indicated by Sin et al. (2010), there is a high correlation between parameters and a large confidence interval for their numerical values. Therefore, a physical meaning cannot be attached to these parameters. This explains why some calibrated parameters are significantly different than those reported by Kadam et al. (2004).

4.2. Steady State average performance simulation scenario

The model inputs are set to the constants given in Table 4, which represent an average performance operating point of the biorefinery. The percentage values from Table 2 for the demonstration plant are transformed into g kg^{-1} in Table 4, given the fact that the fibers have a 25% dry matter. The CO_2 concentration is set to the saturation point, i.e., 0.0011 g kg^{-1} at 50 °C. The simulation runs for 20 h and the stabilized reactor profiles can be observed in Fig. 4 drawn with a solid black line.

Subplot A illustrates the buffer concentrations in mol L^{-1} relevant for pH calculations. The acetic acid concentration, notated as A_T, increases along the reactor due to xylan hydrolysis, as expected. The base concentration B_T, the bicarbonate C_T, the succinic acid S_T and lactic acid L_T are constant throughout the reactor as there is no production or consumption of these species.

Subplot B shows the conversion of cellulose and xylan into cellobiose, glucose and xylose. Cellulose C_{C_S} and xylan C_{X_S} drop along

Table 4
Boundary conditions for a typical demonstration scale process operation used in the average performance simulation scenario.

Input	Description	Value	Unit
F_{FF}	Fiber flow	1.11	kg s^{-1}
F_B	Base flow	0.012	kg s^{-1}
F_E	Enzymes flow	0.025	kg s^{-1}
F_W	Water flow	0.014	kg s^{-1}
F_{FM}	Fiber mash outflow	1.16	kg s^{-1}
C_E	Enzymes concentration	500.0	g kg^{-1}
C_{B_0}	Base concentration	270	g kg^{-1}
Fiber fraction composition			
$C_{C_{S_0}}$	Cellulose concentration	112.5	g kg^{-1}
$C_{X_{S_0}}$	Xylan concentration	20	g kg^{-1}
$C_{L_{S_0}}$	Lignin concentration	80	g kg^{-1}
$C_{A_{T_0}}$	Acetic acid concentration	5.0	g kg^{-1}
$C_{C_{T_0}}$	CO_2 concentration	0.0011	g kg^{-1}
$C_{S_{T_0}}$	Succinic acid concentration	0.4	g kg^{-1}
$C_{L_{T_0}}$	Lactic acid concentration	0.7	g kg^{-1}
C_{G_0}	Glucose concentration	0.5	g kg^{-1}
C_{C_0}	Cellobiose concentration	0.0	g kg^{-1}
C_{X_0}	Xylose concentration	2.5	g kg^{-1}
C_{F_0}	Furfural concentration	1.8	g kg^{-1}
T	Optimal enzymatic activity temperature	50.0	°C

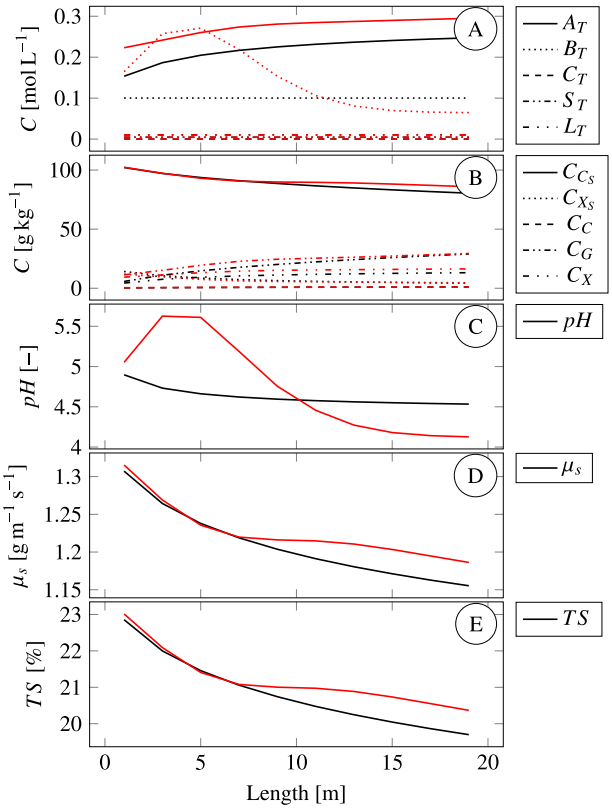

Fig. 4. Reactor profiles from the average performance simulation scenario with constant inputs (drawn with ——) and from time $t = 24$ h when the model is driven by real measurements (drawn with ——). Subplot A shows the buffer concentrations relevant for pH calculations, i.e. acetic acid A_T, base B_T, bicarbonate C_T, succinic acid S_T and lactic acid L_T. Subplot B illustrates cellulose C_{C_S} and xylan C_{X_S} conversion into cellobiose C_C, glucose C_G and xylose C_X. Subplots C–E display the pH, viscosity and solids profiles along the reactor.

the reactor as the liquefaction process progresses. As sugars are formed, the conversion rate decreases due to the inhibition effects of newly created xylose and glucose, which is in accordance with Eqs. (26) and (28). This is why glucose and xylose are formed more rapidly in the first sections of the reactor. Approximately 20% of the entering cellulose is converted into glucose for the given reactor.

Subplots C–E illustrate the pH, viscosity and solids profiles. Acetic acid formation leads to a drop in pH level of 0.4 units along the reactor. This pH deviation affects the optimal performance of the enzymes, which also contributes to a slower conversion. Monitoring the pH profile can tell how optimal a hydrolysis process runs from the point of view of enzymatic activity. The slurry viscosity drops along the reactor by 0.15 units and is a good indicator of how liquid the mixture becomes, i.e. viscous or not. The solids profile shows a reduction of 3% in solids, mainly due to cellulose and xylan liquefaction. The other solids, e.g. lignin, remain in the mixture unchanged. The reactor retention time in this simulation scenario is 7.8 h, which is enough time to create a transportable slurry that can be pumped to subsequent tanks for continuing the hydrolysis process. A typical hydrolysis process requires 140–160 h (Kadam et al., 2004).

4.3. Dynamic simulation scenario with full scale real measurements and uncertainty analysis

Real data from a demonstration scale biorefinery have been recorded throughout 170 h of operation. The model is simulated by using the recorded plant online data including feed composition and flux. A snapshot of the reactor profiles is caught at time $t = 24$ h. The profiles are drawn with red lines in Fig. 4 overlapped with the simulation scenario profiles. In *subplot A* it is seen that real data show a higher concentration of acetic acid in the reactor, which causes a lower enzymatic activity. More base starts to be pumped (shown with dotted red line in *subplot A*) in order to adjust the pH level and an overshoot is recorded as the pH level increases to 5.5 units near the inlet (*subplot C*). Still, the overall enzymatic activity is improved and a conversion close to the theoretical profile found in the previous scenario is recorded. Consequently, the slurry viscosity μ_s and the total solids percentage start to drop as illustrated in *subplots D* and *C*. Conversion still occurs even though the hydrolysis process runs suboptimal from the point of view of the enzymatic activity.

The top plot in Fig. 5 illustrates the outstream pH prediction, which fits relatively well with the sensor data. There are several reasons that can cause estimation differences: the online NIR analyzer of pretreated fibers is the major source of inaccuracies regarding the relevant buffer concentrations for pH calculations; the flow measurement of fibers also presents drifts and offsets leading to an erroneous calculation of the incoming acids affecting the estimated pH; estimation errors could also occur due to unknown buffers created in the pretreatment process; a last source of errors could be the pH sensor itself, which can have offsets if the cleaning procedure is not followed properly. For example, it is suspected that the pH sensor has an offset error between 90 h and 140 h because the pH measurement drops below 4 units, which is unlikely to happen in reality. Also, the measurements in this time frame are very noisy compared to the non-faulty case and there is a sudden change in the measurement at approximately $t = 140$ h, which can happen only if the sensor has been cleaned at that specific time.

The pH estimation is more dynamic than the measurement, presenting several spikes when the reactor holdup is changing but stays most of the time within 4–5 pH units with an error below 1 unit, which is considered satisfactory knowing the complexity of the chemical species in the slurry.

The bottom plot from Fig. 5 shows the calculated fiber mash viscosity or μ_S. The values indicate a rather liquid mixture that can be pumped further to subsequent tanks. Typically, the viscosity near the tank inlet is above 1.3 units as indicated in *subplot D* from Fig. 4.

Throughout the operation of the biorefinery, samples were grabbed at the outflow every 6 h. These samples were then analyzed with an HPLC device in the laboratory in order to observe the sugar concentrations. Fig. 6 displays glucose, xylose and cellobiose concentrations from the outstream, i.e. measurements with bullet points and predicted values with solid lines. The gray area indicates the 5th–95th percentile interval obtained after running the Monte Carlo simulations with LHS sampling on kinetics and feed parameters as explained in the methodology. The sugar formation fits fairly well within the 5th–95th percentile. Some differences appear when the reactor holdup changes, which can be due to non trivial mixing effects that are not captured well by the 1-D transport equation.

The SRC coefficients are summarized in Table 5, where all kinetic and feed parameters are ranked with respect to the β_i coefficients. The ranking shows that most of the model output uncertainty is explained by the uncertainties in the kinetic parameters. The contribution of the feed composition uncertainties to the model output are much smaller. This makes sense because the hydrolysis tank has a large volume and acts as a buffer to feed variations while the intrinsic uncertainties of the enzyme kinetics affect the degree of conversion of cellulose, hence the predicted performance of the reactor.

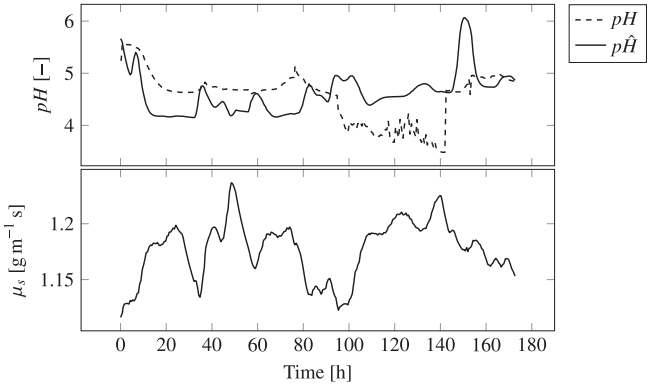

Fig. 5. pH and viscosity calculations for the reactor outstream. The top plot compares the calculated pH, i.e. $p\hat{H}$, against the real measurement pH. The bottom plot shows the calculated slurry viscosity.

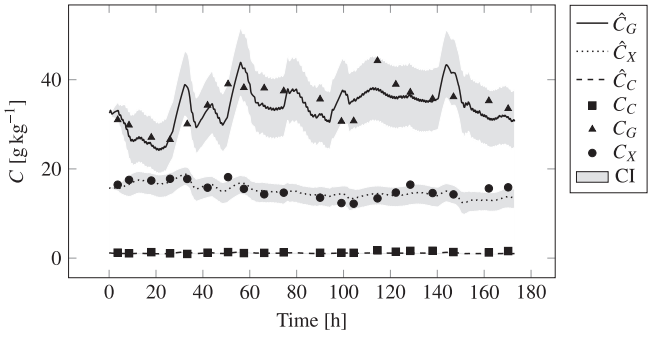

Fig. 6. Cellobiose, glucose and xylose formation during the enzymatic hydrolysis process. $\hat{C}_C, \hat{G}_C$ and $\hat{X}_C$ are estimated concentrations of cellobiose, glucose and xylose as predicted by the model while C_C, C_G and C_X are sugar concentrations measured by the HPLC equipment. The gray area (CI) shows the 5th–95th percentile of model predictions considering kinetics and feed uncertainties.

Table 5
Parameter ranking with respect to the SRC coefficients: β_{i_C}, β_{i_G} and β_{i_X} are the standardized regression coefficients for cellobiose, glucose and xylose, respectively.

Rank	Parameter	β_{i_C}	Parameter	β_{i_G}	Parameter	β_{i_X}
1	K_1	1.00	I_{G_2}	1.00	I_{G_2}	0.99
2	I_{G_2}	1.00	I_{X_1}	0.99	K_2	0.95
3	I_{G_1}	1.00	K_2	0.99	K_4	0.68
4	I_{C_1}	1.00	I_{G_1}	0.71	I_{G_4}	0.56
5	I_{X_2}	0.65	E_{M_1}	0.71	F_μ	0.31
6	I_{X_1}	0.41	K_1	0.59	E_{M_1}	0.21
7	K_2	0.27	K_4	0.42	I_{X_1}	0.17
8	E_{M_1}	0.24	I_{C_1}	0.33	K_1	0.16
9	K_4	0.23	I_{G_4}	0.27	I_{X_2}	0.08
10	F_μ	0.17	K_{A_1}	0.21	I_{C_1}	0.06
11	K_{A_1}	0.14	F_μ	0.17	K_{A_1}	0.06
12	I_{G_4}	0.08	F_σ	0.05	I_{G_1}	0.03
13	F_σ	0.00	I_{X_2}	0.03	F_σ	0.03

4.4. Perspectives

The presented model of the hydrolysis reactor is a promising tool for simulation based process studies. One example is monitoring using soft sensors. A soft sensor is an application of the formulated dynamic model that is able to provide information about variables of interest that cannot be directly measured or, for which, sensors are missing. In the hydrolysis reactor case, the pH and viscosity profiles cannot be directly measured but can be constructed by the model. These profiles constitute important monitoring tools for assessing the performance of the enzymatic process.

Another application deals with fault diagnosis. Algorithms can be developed that investigate the differences between the model predictions and the real measurements in order to detect and isolate any faults that drifts the process from running normally. For example, it could be automatically found when the pH sensor would need to be cleaned.

Process control could also be performed with the current model. Advanced observer based control strategies can be derived for controlling the pH level, the glucose formation or enzyme dosage. Other possibilities will be investigated in the future.

While many promising applications are possible, one area of improvement is in the calculation of viscosity and its comparison to real-measurements which is subject to ongoing work. Additionally, given how uncertainty in model parameters affects the prediction quality, robust techniques can be used for both control and optimization purposes. These are subject to ongoing work.

5. Conclusions

This investigation formulated and analyzed an enzymatic hydrolysis dynamic model for demonstration scale processes. The model consists of a transport module, coupled viscosity estimation and dynamic pH prediction, and pH dependent enzymatic kinetics. The uncertainty analysis performed on the model predictions indicated an acceptable variance that matches well with the measured variability of glucose, xylose and cellobiose of long term plant data. This indicates the quality and reliability of the model as a valuable tool for monitoring, diagnosis and control design.

Acknowledgements

The close collaboration and very helpful suggestions and detailed comments received from Dr. Jakob M. Jensen and Eng. Michael Elleskov on this research are gratefully acknowledged.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.biortech.2013.10.029.

References

Albert, A., Serjeant, E.P., 1962. Ionization Constants of Acids and Bases: A Laboratory Manual. Methuen, London.
Ballesteros, I., Negro, J., Oliva, J.M., Cabañas, A., Manzanares, P., Ballesteros, M., 2006. Ethanol production from steam-explosion pretreated wheat straw. Appl. Biochem. Biotechnol. 129–132, 496–508.
Bird, R.B., Stewart, W.E., Lightfoot, E.N., 2006. Transport Phenomena, Revised 2nd Edition. John Wiley & Sons.
Brun, R., Reichert, P., Künsch, H.R., 2001. Practical identifiability analysis of large environmental simulation models. Water Resour. Res. 37, 1015–1030.
Cantarella, M., Cantarella, L., Gallifuoco, A., Spera, A., Alfani, F., 2004. Effect of inhibitors released during steam-explosion treatment of poplar wood on subsequent enzymatic hydrolysis and ssf. Biotechnol. Prog. 20, 200–206.
Chiaramonti, D., Prussi, M., Ferrero, S., Oriani, L., Ottonello, P., Torre, P., Cherchi, F., 2012. Review of pretreatment processes for lignocellulosic ethanol production, and development of an innovative method. Biomass Bioenerg. 46, 25–35.
Converti, A., Zilli, M., Arni, S., Felice, R.D., Borghi, M.D., 1999. Estimation of viscosity of highly viscous fermentation media containing one or more solutes. Biochem. Eng. J. 4, 81–85.
Datta, R., Maher, M.A., Jonesc, C., Brinkerd, R.W., 2011. Ethanol – the primary renewable liquid fuel. J. Chem. Technol. Biotechnol. 86, 473–480.
Egeland, O., Gravdahl, J.T., 2002. Modeling and Simulation for Automatic Control. Marine Cybernetics.
Glasstone, S., Laidler, K.J., Eyring, H., 1941. The Theory of Rate Processes: The Kinetics of Chemical Reactions, Viscosity, Diffusion and Electrochemical Phenomena. McGraw Hill.
Helton, J., Davis, F., 2003. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. Reliab. Eng. Syst. Safe. 81, 23–69.
Hodge, D.B., Karim, M.N., Schell, D.J., McMillan, J.D., 2009. Model-based fed-batch for high-solids enzymatic cellulose hydrolysis. Appl. Biochem. Biotechnol. 152, 88–107.
Hosseini, S.A., Shah, N., 2011a. Enzymatic hydrolysis of cellulose part ii: Population balance modelling of hydrolysis by exoglucanase and universal kinetic model. Biomass Bioenerg. 35, 3820–3840.
Hosseini, S.A., Shah, N., 2011b. Modelling enzymatic hydrolysis of cellulose part i: Population balance modelling of hydrolysis by endoglucanase. Biomass Bioenerg. 35, 3841–3848.
Jørgensen, H., Vibe-Pedersen, J., Larsen, J., Felby, C., 2007. Liquefaction of lignocellulose at high-solids concentrations. Biotechnol. Bioeng. 96 (5), 862–870.
Kadam, K.L., Rydholm, E.C., McMillan, J.D., 2004. Development and validation of a kinetic model for enzymatic saccharification of lignocellulosic biomass. Biotechnol. Prog. 20, 698–705.
Khan, M.A., 2010. Hydrolysis of Hemicellulose by Commercial Enzyme Mixtures. Master's thesis. Luleå University of Technology. http://epubl.ltu.se/1402-1552/2010/040/LTU-DUPP-10040-SE.pdf.
Larsen, J., Haven, M.Ø., Thirup, L., 2012. Inbicon makes lignocellulosic ethanol a commercial reality. Biomass Bioenerg. 46, 36–45.
Larsen, J., Petersen, M.Ø., Thirup, L., Li, H.W., Iversen, F.K., 2008. The ibus process – lignocellulosic bioethanol close to a commercial reality. Chem. Eng. Technol. 31, 765–772.
McAvoy, T.J., Hsu, E., Lowenthal, S., 1972. Dynamics of ph in controlled stirred tank reactor. Ind. Eng. Chem. Process Des. Dev. 11, 68–70.

Morales-Rodriguez, R., Meyer, A.S., Gernaey, K.V., Sin, G., 2011. Dynamic model-based evaluation of process configurations for integrated operation of hydrolysis and co-fermentation for bioethanol production from lignocellulose. Bioresour. Technol. 102, 1174–1184.

Morales-Rodriguez, R., Meyer, A.S., Gernaey, K.V., Sin, G., 2012. A framework for model-based optimization of bioprocesses under uncertainty: lignocellulosic ethanol production case. Comput. Chem. Eng. 42, 115–129.

Naik, S., Goud, V.V., Rout, P.K., Dalai, A.K., 2010. Production of first and second generation biofuels: A comprehensive review. Renew. Sust. Energ. Rev. 14, 578–597.

Prunescu, R.M., Blanke, M., Sin, G., 2012. Temperature modelling of the biomass pretreatment process. Proc. 17th Nordic Process Cont. Workshop, 8–17.

Prunescu, R.M., Blanke, M., Sin, G., 2013. Modelling and control of ph in bioethanol enzymatic process. Proc. 2013 Am. Cont. Conf. 1, 1891–1898.

Sin, G., Meyer, A.S., Gernaey, K.V., 2010. Assessing reliability of cellulose hydrolysis models to support biofuel process design – identifiability and uncertainty analysis. Comput. Chem. Eng. 34, 1385–1392.

Sin, G., Vanrolleghem, P.A., 2007. Extensions to modeling aerobic carbon degradation using combined respirometric–titrimetric measurements in view of activated sludge model calibration. Water Res. 41, 3345–3358.

Thomas, D.G., 1965. Transport characteristics of suspension: VIII. A note on the viscosity of newtonian suspensions of uniform spherical particles. J. Colloid Sci. 20, 267–277.

Zhang, Y.H.P., Lynd, L.R., 2004. Toward an aggregated understanding of enzymatic hydrolysis of cellulose: noncomplexed cellulase systems. Biotechnol. Bioeng. 88, 797–824.