Model-based plant-wide optimization of large-scale lignocellulosic bioethanol plants.

Prunescu, Remus Mihail; Blanke, Mogens; Jakobsen, Jon Geest; Sin, Gürkan

Published in:
Biochemical Engineering Journal

Link to article, DOI:
10.1016/j.bej.2017.04.008

Publication date:
2017

Document Version
Early version, also known as pre-print

Model-based plantwide optimization of large scale lignocellulosic bioethanol plants

Remus Mihail Prunescu

Automation and Control Group, Department of Electrical Engineering, Technical University of Denmark, Elektrovej Building 326, 2800, Kgs. Lyngby, Denmark

Mogens Blanke

Automation and Control Group, Department of Electrical Engineering, Technical University of Denmark, Elektrovej Building 326, 2800, Kgs. Lyngby, Denmark

Jon Geest Jakobsen

Process Control and Optimization, DONG Energy Thermal Power, Denmark

Gürkan Sin*

CAPEC-PROCESS, Department of Chemical and Biochemical Engineering, Technical University of Denmark, Søltofts Plads Buildings 227 and 229, 2800, Kgs. Lyngby, Denmark

Abstract

Second generation biorefineries transform lignocellulosic biomass into chemicals with higher added value following a conversion mechanism that consists of: pretreatment, enzymatic hydrolysis, fermentation and purification. The objective of this study is to identify the optimal operational point with respect to maximum economic profit of a large scale biorefinery plant using a systematic model-based plantwide optimization methodology. The following key process parameters are identified as decision variables: pretreatment temperature, enzyme dosage in enzymatic hydrolysis, and yeast loading per batch in fermentation. The plant is treated in an integrated manner taking into account the interactions and trade-offs between the conversion steps. A sensitivity and uncertainty analysis follows at the optimal solution considering both model and feed parameters.

*Corresponding author
Email address: gsi@kt.dtu.dk (Gürkan Sin)

It is found that the optimal point is more sensitive to feedstock composition than to model parameters, and that the optimization supervisory layer as part of a plantwide automation system has the following benefits: (1) increases the economical profit, (2) flattens the objective function allowing a wider range of operation without negative impact on profit, and (3) reduces considerably the uncertainty on profit.

Keywords: Second generation bioethanol plant; Nonlinear model-based optimization; Uncertainty and sensitivity analysis; Steam pretreatment; Enzymatic hydrolysis; C5 and C6 co-fermentation.

Nomenclature

β_k	The β coefficient in global sensitivity analysis.
δ_k	Non-dimeansional local differential sensitivity measure of cost function c with respect to parameter θ_k.
$\dot{x}$	Vector of state derivatives used in dynamic modelling.
$\dot{x}_f$	Vector of state derivatives used in fermentation dynamic model.
$\mathbf{R_F}$	Correlation matrix for fermentation model parameters.
$\mathbf{R_L}$	Correlation matrix for liquefaction model parameters.
$\mathbf{R_P}$	Correlation matrix for pretreatment model parameters.
$\mathbf{R}$	Correlation matrix for the entire integrated process.
σ_y	Standard deviation of the objective function.
σ_{θ_k}	Standard deviation of parameter θ_k.
θ_k	Model parameter k. See Table 1 from the supplimentary material for a full list of model parameters.

$c(x, u)$ Cost function as a relation of model states x and decision variables u. [$\mathrm{unitcost}$].

C_b Feedstock composition in [$\mathrm{g/kg}$].

c_f Cost of fermentation.

c_o Value of cost function at the optimal solution.

C_{AC_S} Acetyls concentration [$\mathrm{g/kg}$].

C_{C_S} Cellulose concentration [$\mathrm{g/kg}$].

c_{eh} Cost of enzymatic hydrolysis.

C_{Eth} Concentration of ethanol [$\mathrm{g/kg}$].

c_{ss_k} Cost value in steady-state when varying parameter k.

E_{HMF} 5-HMF activation energy.

$f(x, u)$ Nonlinear process model of states x and inputs u formulated as equality constraints.

F_b Feedstock flow rate [$\mathrm{kg/h}$].

F_e Enzyme dosage [$\mathrm{kg/h}$].

F_s Steam flow rate [$\mathrm{kg/h}$].

$g(x, u)$ Inequality constraints used as ranges for decision variables.

$h(x_f, u_f)$ Dynamic model for C5-C6 co-fermentation.

K_2 Cellulose to glucose reaction constant.

M_y Yeast seed [kg].

M_{Eth} Mass of ethanol [kg].

P_b Feedstock price [$\mathrm{unitcost/(kg/h)}$].

P_e Enzyme price [$\mathrm{unitcost/(kg/h)}$].

P_s Steam price [$\mathrm{unitcost/(kg/h)}$].

P_y Yeast price [$\mathrm{unitcost/kg}$].

P_{MP_X} Ethanol inhibition on xylose uptake.

$q_{Max_{Ac}}$ Maximum acetate uptake rate.

R_B Severity factor dependency.

t_f Final time in fermentation [h].

T_{tr} Thermal reactor temperature [$^{\circ}\mathrm{C}$].

u Vector with all decision variables.

u_f Input variables in fermentation.

x_f Process states in fermentation.

Y_{Cell_G} Biomass growth on glucose.

Y_{Eth_G} Ethanol production from glucose uptake.

Y_{Eth_X} Ethanol production from xylose uptake.

z_i Initial guess for the optimization problem.

z_o Optimal solution.

1. Introduction

Second generation lignocellulosic biorefineries reached commercial reality in 2012 [1], and several large scale plants are in operation nowadays including Beta Renewables, Abengoa Bioenergy, GranBio and POET-DSM [2]. Most biorefineries produce bioethanol, but the drop in oil price reduced the demand on the biofuel. However, plant upgrades for chemicals with higher-added values are pursued making biorefineries still competitive in an oil dependent environment [3].

This study deals with optimizing the daily operation of a large scale second generation biorefinery with a well established conversion route for bioethanol production. The focus is not a techno-economic assessment of alternative biomass conversion technologies but rather assumes a plant has already been built and is in operation.

The latest developments in biorefinery technology show that integrating the facility with a nearby power plant following the Integrated Biomass Utilization System (IBUS) [4] has a major impact on cost efficiency. E.g. the Inbicon plant is integrated with Asnæsværket situated in Kalundborg Denmark and they are both owned by the same company DONG Energy. The symbiosis between the biorefinery and the power plant allows the exchange of by-products for consumables, e.g. lignin bio-pellets for steam.

Modeling and simulation are used in this study as enabling technology to analyze plant performance as basis for an overall optimization. The objective of the optimization problem is to maximize the plant economical profit, considering prices for the most important consumables and end products of the process: biomass, enzymes, yeast and ethanol.

The conversion route from lignocellulosic material to products with higher added value consists of: pretreatment, enzymatic hydrolysis, fermentation, and purification [1, 4]. Lignocellulosic biomass contains cellulose, hemicellulose (xylan and arabinan), lignin, ash, and other residues [5]. The scope of the pretreatment process is to open the biomatrix, relocate lignin and partially hydrolyze the hemicellulose such that cellulose would become more accessible to the downstream process of enzymatic hydrolysis [6]. During pretreatment, inhibitors such as organic acids, furfural, and 5-Hydroxymethylfurfural (5-HMF) are also created due to sugar degradation. Organic acids change the pH of medium, but can be automatically neutralized by a pH controller for ensuring optimal enzymatic conditions [7]. Furfural, 5-HMF, and acetate are fermentation inhibitors [8], while the remaining hemicellulose fraction leads to xylooligomers and xylose formation in the enzymatic hydrolysis process, which strongly inhibit the enzymatic activity [9].

There are trade-offs between the conversion steps. Too little biomass pretreatment would reduce the exposed cellulose to enzymes, and also increases the amount of hemicellulose for enzymatic hydrolysis, which would eventually decrease the glucose yield due to xylose and xylooligomers inhibition. On the other hand, too much biomass pretreatment would increase the amount of fermentation inhibitors leading to a lower ethanol yield.

Most existing studies focus on operational optimization conducting small scale experiments in the laboratory for finding the best pretreatment conditions such that ethanol yield is maximized [10–13]. The traditional focus is on one unit at a time (pretreatment versus enzymatic hydrolysis versus fermentation) but the entire process is rarely considered although the biomass conversion steps are inherently dependent and integrated. The single step methods are suboptimal from an economic point of view as they do not focus on overall process economics. Furthermore, in existing studies, the enzymatic hydrolysis and fermentation processes are usually conducted following a fixed recipe, i.e. no correction action or feedback is taken to counteract the effects of inhibitors. For example, one could increase the enzyme dosage when xylooligomers and xylose inhibit glucose production, or adjust the yeast seed in fermentation to compensate for inhibitors.

Therefore the focus of this paper is on systematic methods and tools to facilitate the further process optimization and daily operation of second generation bioethanol plants. The paper shows how overall optimization can be achieved and how sensitivity and uncertainty can be assessed with respect to feedstock composition and kinetic parameters. A Monte Carlo technique with Latin Hypercube Sampling and correlation control is used for the uncertainty analysis following the methodology from [14, 15].

This paper is structured as follows: the methods section revises the methodology for building the optimization layer for plantwide operation, along with the theoretical part of the sensitivity and uncertainty analysis. The results and discussion follow where the profit curve, costs, and optimal solutions are presented along with their uncertainty bounds. The paper concludes with a summary of all

important findings.

2. Methods

2.1. Second Generation Bioethanol Plant

Figure 1 illustrates a generic large scale second generation biorefinery concept for bioethanol production. The pretreatment process consists of a continuous thermal reactor and a separation press, which were modeled and analyzed in [16, 17]. The thermal reactor is equipped with temperature control for adjusting the reaction temperature T_{tr} [18]. When hemicellulose is hydrolyzed, it produces xylose and arabinose (C5 sugars). After separation, the liquid part containing the C5 sugars is directly pumped into fermentation reactors, bypassing the enzymatic hydrolysis reactors. Cellulose can also be degraded in the pretreatment process, but the produced glucose (C6 sugar) is not lost as it is added to fermentation along with the C5 sugars from the liquid fraction.

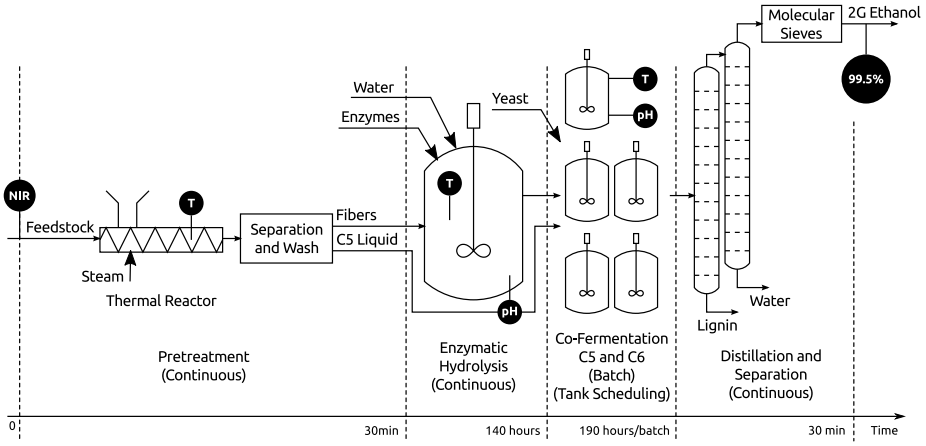


Figure 1: Biorefinery diagram with assumed instrumentation. Pretreatment, enzymatic hydrolysis, and purification are continuous processes, while fermentation occurs in scheduled batch reactors. Feedstock composition is assumed to be known, and can be measured in reality with NIR equipment.

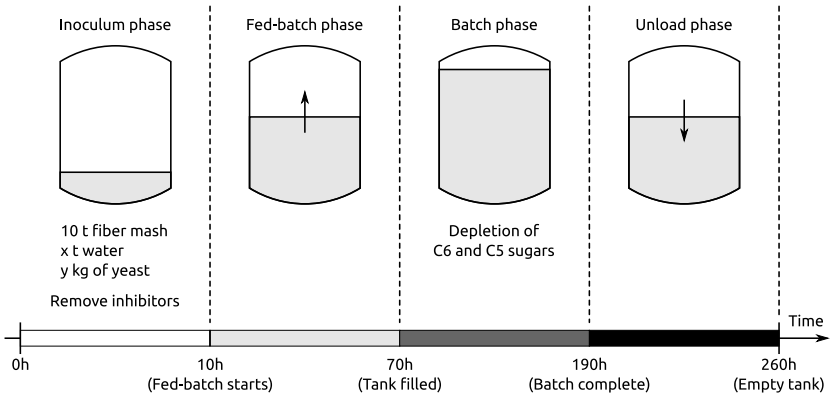
The enzymatic hydrolysis process was thoroughly described and analyzed in [19]. It runs at a high dry matter content in a continuous mode and consists of a series of hydrolysis tanks. The first reactor is described in [20] followed

by conventional continuous stirred tank reactors in order to meet the necessary hydrolysis retention time of $140\,\mathrm{h}$. The tanks are equipped with pH and temperature controllers in order to keep optimal conditions for the enzymatic activity: e.g. $pH = 5$, and $T = 50\,^{\circ}\mathrm{C}$ [21]. Enzymes are added by a pump from a storage tank. The enzyme dosage F_e can be adjusted accordingly and constitutes a degree of freedom in the optimization problem.

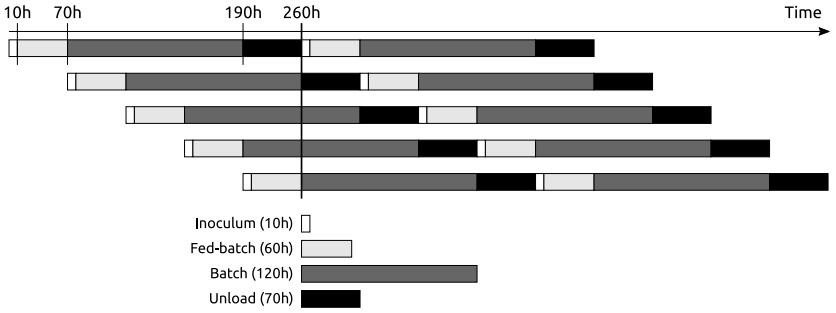
The fermentation process runs in scheduled batch mode in a number of tanks with a maximum holdup of $250\,\mathrm{m}^3$. It is assumed that the size of the reactors has been optimally designed before construction to ensure high titer of ethanol. The fermentation tanks have pH and temperature controllers. Each batch consists of an inoculum phase, the fed-batch phase (which cannot be neglected because it takes $60\,\mathrm{h}$ to fill the tank, time when reactions already take place), the batch stage, and an unload step. These stages are illustrated in Figure 2a. In the inoculum phase, $10\,\mathrm{t}$ of hydrolyzed fibers rich in glucose are mixed with M_y kilograms of yeast and diluted with water. The amount of yeast addition M_y is one of the operation parameters. The fed-batch phase starts as soon as the inhibitors are removed, after about $10\,\mathrm{h}$. The fermentation tank is filled up to $220\,\mathrm{t}$ with a constant feed rate calculated as the sum between the enzymatic hydrolysis outflow rate and the C5 liquid from the pretreatment process. Once the tank is filled, the batch phase begins where the C5 and C6 sugars are slowly depleted. The batch stage has a fixed duration set to $120\,\mathrm{h}$.

A large scale biorefinery has several fermentation reactors running in parallel following a certain scheduling algorithm. Figure 2b shows the scheduling algorithm for 5 reactors such that the overall fermentation inflow and outflow rates have minimum interruptions. This is achieved by aligning in series the fed-batch phases from all tanks, and by synchronizing the unload stages.

The distillation and purification phase separates lignin and water from ethanol. Lignin is recovered as bio-pellets in a evaporation unit as a refinery by-product. The lignin bio-pellets are sent to a nearby power plant where they are co-burnt with coal for steam generation. Lignin bio-pellets are very important for the overall process economics because they contain enough energy



(a) Fermentation process consisting of 4 phases: inoculum ($10\,\mathrm{h}$), fed-batch ($60\,\mathrm{h}$), batch ($120\,\mathrm{h}$), and unload ($70\,\mathrm{h}$).



(b) An example of a fermentation process with 5 scheduled reactors. The reactors are scheduled such that the liquefied fiber inflow and ethanol outflow stay constant with minimum interrupts.

Figure 2: Fermentation process: sequential operation and scheduling.

to produce the required steam for ethanol recovery [22]. The exchange of bio-pellets for steam illustrates the symbiosis between the biorefinery and the nearby power plant following the Integrated Biomass Utilization System (IBUS) [4].

Bioethanol is the main product, which achieves a high concentration of $99.5\,\%$ with the help of several molecular sieves.

The amount of total solids in each conversion step is measured and controlled. This is a key performance feature that enables constant residence times in the reactors and flattens distillation steam consumption.

One can change either the reaction time by modifying the retention time of each individual process, or adjusting the pretreatment temperature, enzyme dosage, and yeast seed to maximize ethanol yield. Since most second generation plants are new, the focus of this study is to maximize the daily operation efficiency of the plant assuming a constant throughput in the refinery, also defined as Mode I in [23]. This constraint translates to a fixed pretreatment time, i.e. $15\,\mathrm{min}$, a constant enzymatic hydrolysis time of $140\,\mathrm{h}$, and a fermentation time of $190\,\mathrm{h}$ per batch for the demonstration scale plant studied here. The degrees of freedom then become: the thermal reactor temperature T_{tr}, the enzyme dosage F_e, and yeast seed M_y, which are the key process parameters to use in the overall optimization.

For commercial scale plants, once the efficiency has been maximized, the optimization problem can be extended to maximize production or search for a maximum throughput by adding the raw material inflow as a degree of freedom in the optimization problem. Maximizing the throughput, defined as Mode II in [23], is not in focus in this study.

2.2. The Optimization Layer

This study aims to develop an optimization method for maximizing the daily operational efficiency that would be implemented as a supervisory layer at a large scale facility. Figure 3 shows the role of the optimization layer, and how it interacts with the system identification layer, the control system, and the real plant. The model identification layer utilizes real measurements to calibrate

the plant model such that predictions become more accurate. The control layer translates the optimal setpoints into actuator commands to ensure reference tracking for the key process states or variables. For pretreatment, temperature regulation is the key variable [18]. For enzyme dosage, the flow of enzymes is controlled. Yeast seed control requires a mass estimator and control of added yeast. Common practice for a new plant is to design the control system to keep a constant throughput [23].

The dynamic models used in this study were validated against data that were collected at the Inbicon demonstration scale plant for a throughput of $1\,\mathrm{t/h}$ of raw biomass. This flow rate was chosen in order to minimize the impact of pretreatment disturbances on fiber composition. At higher throughputs vertical temperature gradients appear in the thermal reactor that create layers of different biomass composition [16].

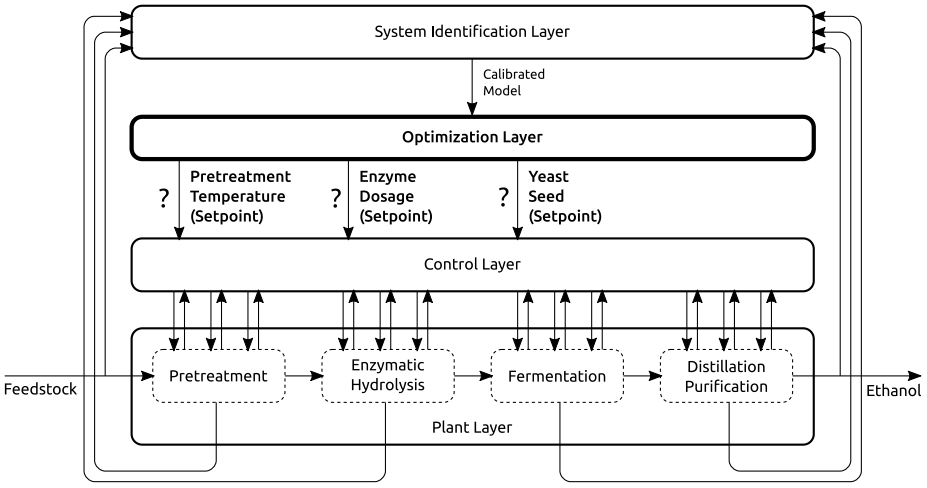


Figure 3: Block diagram showing the interaction between the optimization layer designed in this study and the real plant. The optimization layer calculates setpoints for pretreatment temperature, enzyme dosage and yeast seed. All models are calibrated by the system identification layer based on plant measurements.

The optimization layer is solved or updated either when the underlying models are recalibrated to fit new acquired data, when feedstock composition changes (e.g. due to different biomass type or variability in feedstock content,

which can be measured either offline or online with NIR equipment), or when prices change (e.g. ethanol price increases, enzymes price decreases, yeast can be grown at a lower price etc.). Model calibration and feedstock composition can change on a daily basis while prices are set through contracts with suppliers and remain constant over a much longer period.

The solution of the optimization problem provides setpoints for the pretreatment temperature, enzyme dosage, and yeast seed. The system constraints are formulated based on validated large scale models for: hydrothermal pretreatment with steam [17], enzymatic hydrolysis [19], and C5 and C6 co-fermentation [24]. The optimization problem is formulated as a constrained nonlinear programming problem (NLP) with the following generic formulation [25]:

$$\begin{aligned} \max_{u} \quad & c(x,u) \\ \text{subject to} \quad & f(x,u) = 0 \\ & g(x,u) \leq 0 \end{aligned} \tag{1}$$

where u is a vector of the decision variables or degrees of freedom, x represents the process variables, and $c(x,u)$ is the nonlinear objective function. $f(x,u) = 0$ and $g(x,u) \leq 0$ are equality and inequality nonlinear constraints as functions of process and decision variables. The solution of optimization problem (1) is found by a nonlinear programming solver (e.g. the optimization toolbox in Matlab contains function `fmincon`, which finds the minimum of a constrained nonlinear multivariate function).

2.3. Mathematical Models

The optimization layer uses a dynamic biorefinery simulator to calculate the stabilized or steady-state outputs for the continuous processes of pretreatment and enzymatic hydrolysis, and the final states at the end of the batch fermentation. The mathematical models are complex and nonlinear. Finding an analytical solution for steady-states might not be feasible. An alternative is to run a sufficiently long simulation until all outputs reach steady state. The fermentation model is a batch process and the outputs are collected after running a dynamic

simulation for $190\,\mathrm{h}$, i.e. the end of the batch phase as shown in Figure 2a, which is a fixed amount of time by design due to setting a constant throughput of material in the biorefinery. The pretreatment and enzymatic hydrolysis models have already been published in [17] and [19]. The fermentation tank is modeled as a continuous stirred tank reactor (CSTR) with reaction kinetics derived from [24, 26]. The equations of all models used to simulate the integrated process are provided in the supplementary material.

The plant simulator is implemented in modules, one for each conversion step, i.e. pretreatment, enzymatic hydrolysis and fermentation. Each module has inputs, internal states, outputs and a set of parameters. The inputs and the outputs are data structures containing information on flow rate (in $\mathrm{kg/h}$), composition (in $\mathrm{g/kg}$) and enthalpy (in $\mathrm{kJ/kg}$). E.g. the pretreatment block has 1 biomass input (i.e. the feedstock), which has a feed rate set to 1000 kg/h with the composition from Table 2, and 1 steam input. The thermal reactor is temperature controlled and the steam valve is adjusted accordingly to reach the desired pretreatment temperature. The pretreatment block has 2 outputs: 1 of fibers with high dry matter and 1 liquid stream rich in C5 sugars. The fibers output from the pretreatment block is connected to the input of the enzymatic hydrolysis block, while the output with C5 sugars is connected to one of the inputs of the fermentation block. The inputs and outputs of each block and their interconnection are represented with arrows in Figure 1.

Table 1 offers a summary of the integrated model complexity. The overall model accounts for 96 kinetic parameters, 580 states, 10 inputs and 25 outputs. The table also offers a split of the modular model based on the conversion step. The high number of states in pretreatment and enzymatic hydrolysis is due to the computational fluid dynamics tools (the convection equation discretized in space) used for modeling the thermal reactor and the first enzymatic hydrolysis tank. Nominal values for kinetic and feed parameters are given in Table 1 in the supplementary material.

Table 1: Model summary: number of parameters, states, inputs and outputs. Half of the pretreatment outputs (the pretreated fibers) are directed to the enzymatic hydrolysis process, while the other half (the C5 liquid) is connected to fermentation. The outputs from the enzymatic hydrolysis are connected to fermentation.

Model	Parameters	States	Inputs	Outputs
Pretreatment	17	298	10	36
Enzymatic Hydrolysis	46	257	18	19
Fermentation (1 tank)	33	25	37	25
Total	96	580	10	25

2.4. Plantwide Optimization Methodology

The methodology steps for finding the optimal operational point of a plant are extended from [25]:

1. Select the objective or cost function;
2. Identify the decision variables;
3. Formulate process model constraints and set bounds for decision variables;
4. Formulate and solve the NLP optimization problem;
5. Sensitivity and uncertainty analysis of the optimal solution.

The optimal solution is analyzed from a sensitivity point of view using similar tools as in [14, 15, 17]. Mathematical models that describe complex systems are often over-parametrized. The sensitivity analysis quantifies the relation between the cost function and model parameters when the system runs at the optimal point. The aim is to rank all model parameters by their significance with respect to the profit value at the nominal operational point. Also a subset of relevant parameters can be extracted for calculating the uncertainty bounds.

A non-dimensional measure of local sensitivity suitable for steady-state signals is the differential sensitivity measure defined in [27] and [28]:

$$\delta_k = \frac{\partial c_k}{\partial \theta_k} \frac{\theta_k}{c_{ss_k}} \tag{2}$$

where $\partial c_k / \partial \theta_k$ is the variation in profit with respect to a model parameter, and is calculated based on finite differences. θ_k is the k^{th} parameter, and c_{ss_k} is the profit or the value of the cost function in steady-state. All model parameters are sorted in descending order with respect to δ_k, and a subset is created with all parameters that have δ_k above a threshold. The reduced subset of model parameters reduces model complexity and is then used in the uncertainty analysis.

The propagation of uncertainty is analyzed with a Monte Carlo procedure as described in [14]:

1. Define input uncertainty;
2. Parameter sampling;
3. Monte Carlo simulations;
4. Output uncertainty.

The input uncertainty is defined with standard deviations and correlation matrices obtained from previous studies. Dealing with many parameters implies a large number of combinations of parameter values with high correlation between them. In order to reduce the number of parameter samples but preserve statistical meaning, a Latin Hypercube Sampling (LHS) technique with correlation control is utilized [29]. LHS generates less samples of parameters but is made statistically plausible with the help of a distribution function, standard deviation, and correlation matrix. For each set of samples, a simulation is then run and the output is collected. After all Monte-Carlo simulations are performed, enough output information is obtained to statistically compute the median and the 5th-95th percentile confidence intervals.

A global sensitivity analysis supplements the Monte-Carlo simulations results. The statistical distribution of both the model parameters and the biomass inflow composition are taken into account. The methodology includes two steps [15, 30]: (1) fitting a linear model from each variable of interest (model parameter and biomass composition) to the optimization objective function, which is the biorefinery profit calculated as in Equation (6) and generated through Monte Carlo simulations; (2) calculating the standardized regression coefficients as

a global sensitivity measure, which indicate how much of the output standard deviation is explained by the standard deviation in each input variable.

Step (1) finds coefficients a and b_k from Equation (3) using a least squares method. The linear model fits the objective function while the input parameters are contained in vector θ, i.e. both model parameters θ_R and feedstock composition C_b.

$$y_{reg} = a + \sum_k b_k \theta_k \tag{3}$$

Step (2) uses Equation (4) to calculate the β coefficients, which reflects how much of the output variation is explained by parameters uncertainty:

$$\beta_k = \frac{\sigma_{\theta_k}}{\sigma_y} b_k \tag{4}$$

3. Results and Discussion

This section first presents values of model and feed parameters. The optimization problem is then formulated, it is shown how a solution is found and the properties of the results are discussed. A sensitivity analysis with respect to model kinetic parameters is then presented. A subset of sensitive parameters is identified, and Monte Carlo simulation is employed to quantify the uncertainty of the optimal solution. The costs and profit curves are also computed with uncertainty bounds. Furthermore, a global sensitivity analysis is made in order to identify bottlenecks in model predictions with respect to feed and model parameters. Uncertainty is then embedded in the formulation of a stochastic optimization problem. As a final result, dynamic simulations show the refinery operation at the optimal point.

3.1. Model Initialization

Table 2 shows the feed parameters, i.e. raw biomass inflow rate, composition and initial temperature. The inflow rate is set to $1000\,\mathrm{kg/h}$, the throughput of a demonstration scale plant. The biomass composition resembles wheat straw with an initial dry matter of $89\,\%$ [6].

Table 2: Biorefinery inputs: inflow rate, raw biomass composition, and initial temperature.

	Description	Value	Unit	% of dry matter
1	Inflow rate	1000	kg/h	
2	Cellulose	360	g/kg	40.45
3	Xylan	187	g/kg	21.01
4	Arabinan	23	g/kg	2.58
5	Lignin	200	g/kg	22.47
6	Acetyls	44	g/kg	4.94
7	Ash	26	g/kg	2.92
8	Water	110	g/kg	-
9	Other	50	g/kg	5.63
10	Temperature	15	°C	

Table 1 from the supplementary material indicates the values with units for all 96 model parameters. The table is split into pretreatment, enzymatic hydrolysis and fermentation. The model parameter values are taken from [17] (pretreatment), [19] (enzymatic hydrolysis) and [24] (C5 and C6 co-fermentation).

3.2. The Optimization Problem

The application of the optimization methodology is now highlighted for the case study:

1. Select the objective or cost function:

   The cost function from this study represents the profit for one fermentation batch defined as the difference between ethanol revenue and operating costs related to biomass, steam, enzymes, and yeast:

   $$c(M_{Eth}, F_b, F_s, F_e, M_y) = M_{Eth}(t_f)P_{Eth} - (F_bP_b + F_sP_s + F_eP_e + M_yP_y) \tag{5}$$

   Ethanol revenue is calculated as $M_{Eth}(t_f)P_{Eth}$, i.e. mass of ethanol in kg at the end of the batch phase t_f times its price per kilogram P_{Eth}. M_y is

the mass of yeast added to the fermentation tank in the inoculum phase. The operating costs are defined as flow rate or mass of utility times its price. The refinery consists of two continuous processes, i.e. pretreatment and enzymatic hydrolysis, and a batch process, i.e. fermentation. The weights P_b (cost of biomass), P_s (cost of steam), and P_e (cost of enzymes) are related to the continuous processes, i.e. pretreatment and liquefaction, and are measured in $\mathrm{unitcost}/(\mathrm{kg}/\mathrm{h})$. P_y (cost of yeast) is measured in $\mathrm{unitcost}/\mathrm{kg}$. The overall measuring unit of the cost function becomes the $\mathrm{unitcost}$, which can represent any currency.

Capital costs do not enter the objective function because they are not functions of the decision variables but rather a fixed amount. Capital costs diminish the potential profit by a constant and should be taken into consideration to give a full meaning to the profit solution from the optimization problem but they do not change the optimal solution.

The feedstock flow rate or refinery throughput F_b is kept constant. Product $F_b P_b$ becomes a constant and its derivative with respect to any of the decision variables is 0. Therefore it can be dropped from the cost function. Distillation costs are neglected because the second generation biorefinery is considered to be integrated with a power plant exchanging lignin bio-pellets for steam.

The cost function then becomes:

$$c(M_{Eth}, F_e, M_y) = M_{Eth}(t_f)P_{Eth} - (F_e P_e + M_y P_y) \tag{6}$$

Table 3 shows the weight values used in this optimization study. The weights from Table 3 are not fixed values but rather used as an example for the method. Changing the weights would naturally lead to a different solution, which is a matter of plant operation managers priority. The values depend on procurement contracts of feedstock and enzymes, and the relative importance may change.

2. Identify the decision variables:

   The outcome of the pretreatment process is sensitive to the thermal reactor

Table 3: Cost function weights (prices).

Parameter	Description	Value
P_{Eth}	Ethanol	$5\,\mathrm{unitcost/kg}$
P_e	Enzymes	$25\,\mathrm{unitcost/(kg/h)}$
P_y	Yeast	$50\,\mathrm{unitcost/kg}$

temperature and retention time [11]. When a constant throughput is required, the retention time is constant. The thermal reactor temperature then becomes the only degree of freedom in pretreatment.

The key parameters in enzymatic hydrolysis are: pH, temperature, and concentration of enzymes. The enzymatic activity is a function of pH and temperature, which resemble Gaussian curves with single peaks at pH of $5\,\mathrm{units}$ and temperature $50\,^{\circ}\mathrm{C}$ [19]. Any deviations from the optimal point would reduce the enzymatic efficiency. Control loops keep the pH and temperature close to optimality [7] and it is not indicated to vary these variables. However, the concentration of enzymes can be adjusted by changing the inflow rate of enzymes F_e and constitutes the only degree of freedom in enzymatic hydrolysis for the optimization problem.

The efficiency of the fermentation process is a function of pH, temperature, and yeast seed. The optimal pH level of the GMO yeast is relatively close to that of the enzymes, i.e. 5.5 units. The optimal fermentation temperature is $35\,^{\circ}\mathrm{C}$, which is different from the enzymatic optimal temperature. Controllers keep the pH and temperature conditions at the GMO yeast optimal levels throughout the entire fermentation process. The only degree of freedom considered in fermentation for the optimization problem is the yeast seed M_y in the inoculum phase.

In summary, the decision variables are: the pretreatment temperature T_{tr} defined as the set-point for the thermal reactor temperature controller, the inflow rate of enzymes F_e expressed in $\mathrm{kg/h}$, and the yeast seed M_y in kg

as a set-point for the amount of yeast used to start the inoculum phase:

$$u = [T_{tr} \quad F_e \quad M_y]^\top \tag{7}$$

3. Process model constraints, and bounds for decision variables:
   The dynamic integrated models for pretreatment and enzymatic hydrolysis are detailed in the supplementary material, and are formulated as:

$$\dot{x} = f(x, u) \tag{8}$$

   where $f(x, u)$ is a nonlinear function of states x and inputs u. The steady states are then found as the solution of $\dot{x} = 0$:

$$0 = f(x, u) \tag{9}$$

   Due to the model complexity and the nonlinear nature of $f(x, u)$, an analytical solution to (9) cannot be easily found. As an alternative, the steady states are calculated by running a sufficiently long simulation till all states stabilize.
   The dynamic model for fermentation is described by:

$$\dot{x}_f = h(x_f, u_f) \tag{10}$$

   where $h(x_f, u_f)$ represents a nonlinear complex model of states x_f and inputs u_f. The final states at time $190\,\mathrm{h}$ are found by integrating the model numerically (dynamic simulation):

$$x_f(t_f) = \int_0^{t_f} h(x_f, u_f) dt \tag{11}$$

   where $t_f = 190\,\mathrm{h}$, i.e. the end of the batch phase.
   The decision variables are bounded as follows:

$$\begin{array}{rcccl} 150 & \leq & T_{tr} & \leq & 210\,^\circ\mathrm{C} \\ 10 & \leq & F_e & \leq & 1000\,\mathrm{kg/h} \\ 10 & \leq & M_y & \leq & 1000\,\mathrm{kg} \end{array} \tag{12}$$

   which allows a wide range of operation for searching the optimal point.

4. Formulate and solve the overall NLP optimization problem:

$$\begin{aligned} \max_{T_{tr},F_e,M_y} \quad & M_{Eth}(t_f)P_{Eth} - (F_eP_e + M_yP_y) \\ \text{subject to} \quad & 0 = f(x(t), u(t)) \\ & \dot{x}_f = h(x_f, u_f) \\ 150 \;\leq\; & T_{tr} \;\leq\; 210\,^{\circ}\mathrm{C} \\ 10 \;\leq\; & F_e \;\leq\; 1000\,\mathrm{kg/h} \\ 10 \;\leq\; & M_y \;\leq\; 1000\,\mathrm{kg} \end{aligned} \tag{13}$$

The previous optimization problem is solved numerically by a constrained minimization solver (e.g. *fmincon* from Matlab) set to use a sequential quadratic programming algorithm (SQP) with scaled objective and constraints. The initial solution guess is picked to be feasible, and set to:

$$z_i = \begin{bmatrix} T_{tr} \\ F_e \\ M_y \end{bmatrix} = \begin{bmatrix} 170\,^{\circ}\mathrm{C} \\ 150\,\mathrm{kg/h} \\ 200\,\mathrm{kg} \end{bmatrix} \tag{14}$$

The initial states of all models are set to 0. The solver takes approximately $3\,\mathrm{min}$ to converge on a computer equipped with an Intel i7-5600U CPU. The following optimal point is reached:

$$z_o = \begin{bmatrix} T_{tr} \\ F_e \\ M_y \end{bmatrix} = \begin{bmatrix} 172\,^{\circ}\mathrm{C} \\ 110\,\mathrm{kg/h} \\ 142\,\mathrm{kg} \end{bmatrix} \tag{15}$$

The thermal reactor temperature should be set to $172\,^{\circ}\mathrm{C}$, the enzyme dosage is of about $110\,\mathrm{kg/h}$, and the yeast seed is of $142\,\mathrm{kg}$. This optimal set point gives a profit of:

$$c_o = 7.6714 \times 10^4\,\mathrm{unitprofit} \tag{16}$$

disregarding raw biomass, distillation and capital costs.

In order to gain process insight and to observe how pretreatment conditions affect the downstream processes, an iteration is created through pretreatment

temperatures between $160\,^{\circ}\mathrm{C}$ to $180\,^{\circ}\mathrm{C}$ with a step of $1\,^{\circ}\mathrm{C}$. Each pretreatment temperature changes the fibers and C5 liquid composition. The enzymatic hydrolysis and fermentation processes are then optimized in an integrated manner for each pretreatment temperature:

$$\begin{aligned} \max_{F_e, M_y} \quad & M_{Eth} P_{Eth} - (F_e P_e + M_y P_y) \\ \text{subject to} \quad & 0 = f(x(t), u(t)) \\ & \dot{x}_f = h(x_f, u_f) \\ & 10 \leq F_e \leq 1000\,\mathrm{kg/h} \\ & 10 \leq M_y \leq 1000\,\mathrm{kg} \end{aligned} \tag{17}$$

In this way the pretreatment, liquefaction and fermentation costs, as well as refinery profit can be observed with respect to pretreatment conditions. The same methodology can be applied even if there are recycles between fermentation and liquefaction because these two processes are analyzed in an integrated manner in optimization problem (17). Algorithm 1 shows how to calculate the curves for profit, costs, and optimal solution as functions of pretreatment temperature. z is the optimal solution returned by the optimization problem solver, c_{eh}, c_f are the enzymatic hydrolysis, and fermentation costs, respectively. c is the value of the cost function or the profit. Algorithm 1 is used later in Algorithm 2 to determine the uncertainty bounds on profit, costs, and setpoints with respect to the pretreatment temperature.

3.3. Sensitivity Analysis of Profit Value at the Optimal Point

Figure 4 illustrates the sensitivity analysis of the profit curve with respect to all model parameters calculated as in Equation (2) at the optimal solution (15). A sensitivity threshold separates important parameters from unimportant ones. The sensitivity threshold is typically based on the context of application [30]. E.g. $1\,\%$ as a threshold states that any factors not contributing with $1\,\%$ to the output variance are unimportant while the ones above that are important for further considerations (e.g. parameter estimation, optimization). For this particular application a value of $4.6\,\%$ is chosen. A lower value would increase

Algorithm 1 Calculate optimal costs and profit with respect to pretreatment temperature given a fixed set of model parameters θ and feedstock composition C_b.

function $[z, c_{eh}, c_f, c]$ = COSTS(θ, C_b)

  Set a range of pretreatment temperatures $T_{tr} \leftarrow 160\,^{\circ}\mathrm{C} : 1\,^{\circ}\mathrm{C} : 180\,^{\circ}\mathrm{C}$

  Set initial solution guess to $z_0 \leftarrow [100\,\mathrm{kg/h} \quad 80\,\mathrm{kg}]^{\top}$

  for Each temperature in T_{tr} **do**

    Run pretreatment process at temperature T_{tr_i} and obtain composition of pretreated fibers and C5 liquid.

    $z_i \leftarrow$ Solution of optimization problem (17) given the pretreated fibers composition and C5 liquid from previous step as inputs. Use as initial guess the solution from previous iteration z_{i-1}.

    Calculate mass of ethanol at final fermentation time: $M_{Eth_i} \leftarrow M(t_f) \cdot C_{Eth}(t_f)$ where t_f is the final batch time, $M(t_f)$ is the reactor mass in kg at time t_f, and $C_{Eth}(t_f)$ is the ethanol concentration at time t_f in $\mathrm{g/kg}$.

    Enzyme dosage: $F_{e_i} \leftarrow z_i(1)$.

    Yeast seed: $M_{y_i} \leftarrow z_i(2)$.

    Calculate liquefaction cost: $c_{eh_i} \leftarrow F_{e_i} P_e$.

    Calculate fermentation cost: $c_{f_i} \leftarrow M_{y_i} P_y$.

    Calculate revenue: $r_i \leftarrow M_{Eth_i} P_{Eth}$.

    Calculate profit $c_i \leftarrow r_i - (c_{eh_i} + c_{f_i})$.

  end for

end function

the subset of important parameters but with little or negligible impact on the solution of the optimization problem.

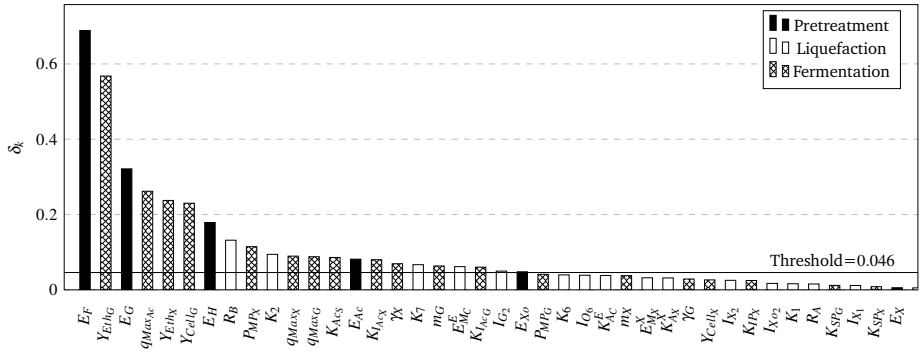


Figure 4: Sensitivity measure δ_k of profit value with respect to model parameters. The model parameters set is reduced to 22 significant parameters out of 96.

The most sensitive parameter is E_F, i.e. the furfural formation activation energy. Furfural is a strong fermentation inhibitor produced during pretreatment, and ethanol yield is directly affected by the amount of furfural. The next sensitive parameter is Y_{Eth_G}, a yield parameter indicating the amount of ethanol in g produced per $1\,\mathrm{g}$ of glucose. E_G or glucose activation energy follows indicating that cellulose degradation in pretreatment impacts the ethanol yield. Three more fermentation parameters with similar sensitivity follow, i.e. maximum acetate uptake parameter $q_{Max_{Ac}}$, cell biomass yield on glucose Y_{Cell_G}, and ethanol yield on xylose Y_{Eth_X}. 5-HMF production during pretreatment has a relatively high sensitivity too as it influences both the glucose yield, by degrading it further, and also by inhibiting ethanol production in fermentation.

The first sensitive enzymatic hydrolysis parameter is R_B, i.e. the severity dependence of the enzymatic activity. R_B shows the importance of biomatrix opening from the pretreatment process as a structural breakdown of the fiber, which affects cellulose accessibility for enzymes. Other important liquefaction parameters are K_2 and K_7, which indicate glucose production rate and enzyme deactivation in time.

The placement of pretreatment parameters such as E_F, E_G, E_H, and E_{Ac}

among the sensitive parameters shows how important pretreatment conditions are for downstream. Fermentation parameters are also numerous among the sensitive parameters. Fermentation with enhanced GMO yeast for bioethanol production is a key process in the biorefinery together with steam pretreatment. Liquefaction parameters have a lower importance because the overall hydrolysis time is long enough to compensate for any parameter uncertainties. The liquefaction process has a pure hydrolysis phase of $140\,\mathrm{h}$ followed by fermentation where enzymes are still active continuing cellulose degradation (simultaneous saccharification and fermentation).

The sensitivity threshold is set at 0.046, which reduces the parameters count to 22 out of 96 showing the importance of the sensitivity analysis in reducing model complexity. These parameters are then used in the following uncertainty analysis.

3.4. Uncertainty Analysis of Costs, Profit and Optimal Solution

The standard deviation and correlation matrix for pretreatment and enzymatic hydrolysis parameters is obtained from [17] and [18, 28], respectively. Table 4 displays a list of all sensitive parameters with their assumed uncertainty as a normal distribution of a mean value and standard deviation. Regarding the model parameters for fermentation, there is no published real data that could be used for uncertainty characterization. Therefore, this study follows techniques from experimental design for system identification, and generates measurements through simulation, i.e. glucose, xylose and ethanol levels, which can be obtained in reality from sample based HPLC readings. Normally distributed measurement noise with $5\,\%$ standard deviation is added, and then a parameter estimation procedure runs on the simulated data for estimating the standard deviation and correlation matrix for these parameters. The results are included in Table 4 where the standard deviation for fermentation parameters is obtained from simulated data as previously described.

The correlation matrix for the sensitive parameters presented in Table 4 is

Table 4: List with all sensitive parameters with their assumed distribution for uncertainty analysis. N(m, σ) stands for normally distributed with mean m and standard deviation σ.

Parameter	Distribution	Reference
Pretreatment		
E_F	N$(327\,255, 285)$	[17]
E_G	N$(335\,616, 249)$	[17]
E_H	N$(299\,999, 2639)$	[17]
R_B	N$(2.915, 0.048)$	This study
E_{Ac}	N$(242\,693, 174)$	[17]
E_{Xo}	N$(298\,010, 98)$	[17]
Enzymatic Hydrolysis		
K_2	N$(0.0065, 0.0001)$	[18, 28]
K_7	N$(2.5 \times 10^{-7}, 0.041\,25 \times 10^{-7})$	[18, 28]
$E^E_{M_C}$	N$(0.016, 0.000\,26)$	[18, 28]
I_{G_2}	N$(0.067, 0.0011)$	[18, 28]
Fermentation		
Y_{Eth_G}	N$(0.47, 6.0744 \times 10^{-3})$	This study
$q_{Ac_{Max}}$	N$(1.23 \times 10^{-5}, 8.2395 \times 10^{-7})$	This study
Y_{Eth_X}	N$(0.4, 1.3127 \times 10^{-2})$	This study
Y_{Cell_G}	N$(0.115, 4.3453 \times 10^{-3})$	This study
P_{MP_X}	N$(100.2, 1.2852)$	This study
q_{Max_X}	N$(0.8 \times 10^{-3}, 1.6115 \times 10^{-5})$	This study
q_{Max_G}	N$(0.3 \times 10^{-3}, 6.7299 \times 10^{-6})$	This study
K_{Ac_S}	N$(2.5, 5.7082 \times 10^{-2})$	This study
$K_{I_{Ac_X}}$	N$(0.2, 8.9583 \times 10^{-3})$	This study
γ_X	N$(0.608, 4.0872 \times 10^{-2})$	This study
m_G	N$(2.6944 \times 10^{-5}, 8.6390 \times 10^{-7})$	This study
$K_{I_{Ac_G}}$	N$(2.74, 4.2621 \times 10^{-2})$	This study

built as a block diagonal matrix with the following structure:

$$\mathbf{R} = \begin{bmatrix} \mathbf{R_P} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{R_L} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R_F} \end{bmatrix} \tag{18}$$

where $\mathbf{R_P}$, $\mathbf{R_L}$ and $\mathbf{R_F}$ are full matrices, and represent the correlation of the sensitive pretreatment, liquefaction and fermentation parameters grouped by refinery step. It is assumed that the refinery steps are not correlated with each other, e.g. pretreatment parameters are not correlated with liquefaction and fermentation parameters. Therefore, $\mathbf{R}$ has a block diagonal shape. The numerical values for $\mathbf{R_P}$, $\mathbf{R_L}$ and $\mathbf{R_F}$ are included in the supplementary material. $\mathbf{R_P}$ and $\mathbf{R_L}$ are obtained from [17] and [28], respectively. The correlation matrix for fermentation parameters $\mathbf{R_F}$ is obtained in this study from simulated data.

Latin Hypercube Sampling (LHS) with correlation control can then be performed for all model parameters. Feedstock composition is sampled assuming uniform distribution with $5\,\%$ variation in composition. $N = 200$ LHS samples are extracted for model and feed parameters. The Monte Carlo simulations are performed by running Algorithm 1 for each set of parameters. The simulation outputs are collected, and the 5th, 50th and 95th percentiles are then calculated for profit, costs, and optimal solution.

The uncertainty analysis is carried out separately for feed, and then for model parameters in order to observe the contribution of each source of uncertainty. A last analysis combines the samples of feed and model parameters to find the overall effect of the uncertainty sources on the outputs.

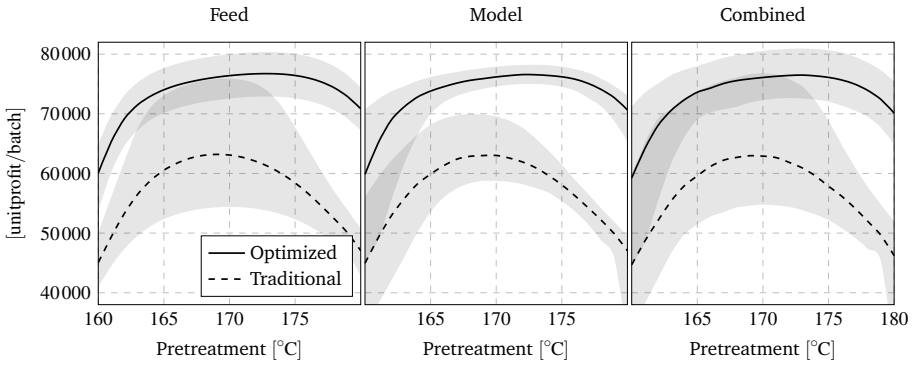
The entire sensitivity and uncertainty analysis for combined model and feed parameters is summarized in Algorithm 2. To separate model and feed parameters step 11 from Algorithm 2 is modified either by keeping C_b or θ constant.

The results of the uncertainty analysis are displayed in Figure 5 and commented below:

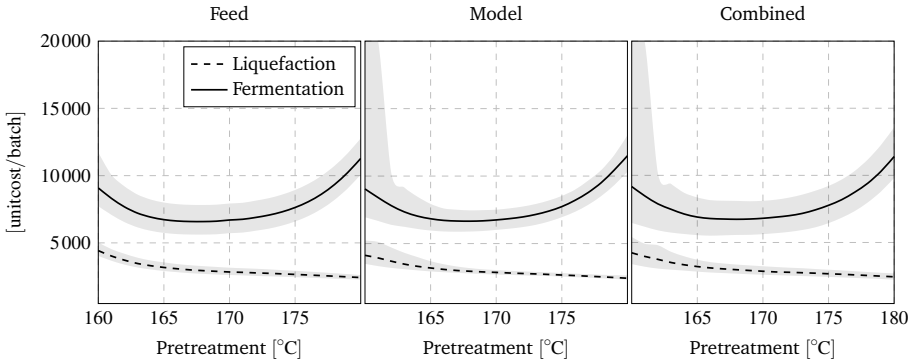
- The profit curve is drawn in Figure 5a, which is used to identify the optimal

Algorithm 2 Sensitivity and uncertainty analysis with combined model and feed parameters.

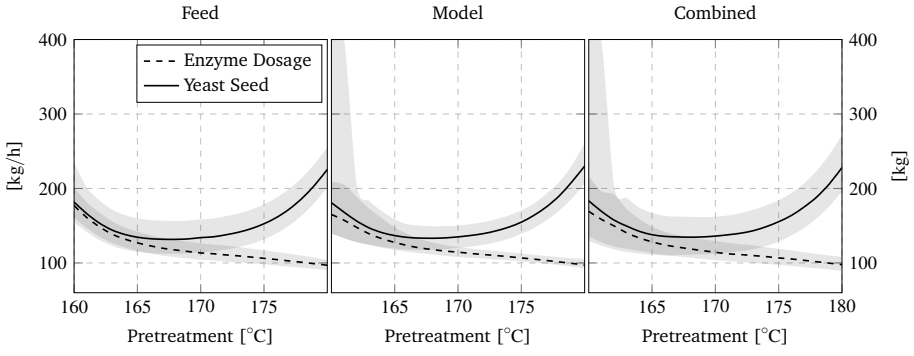
1: Initialize model parameters θ and feed composition C_b.
2: Optimal deterministic solution: $[T_{tr}, F_e, M_y] \leftarrow$ (13).
3: Sensitivity analysis of the cost function in $[T_{tr}, F_e, M_y]$: calculate δ_k as in equation (2).
4: Rank all parameters with respect to δ_k.
5: Select a subset θ_R such that θ_R is above a threshold.
6: Set standard deviations and correlation matrices for θ_R.
7: $\theta \leftarrow$ LHS of θ_R with correlation control to generate N sets of model parameters.
8: Set bounds for feedstock composition.
9: $C_b \leftarrow$ Uniform LHS for feedstock composition to generate N sets of compositions.
10: **for** Each set of model and feed parameters **do**
11: $\quad [z_i, c_{eh_i}, c_{f_i}, c_i] = \text{COSTS}(\theta_i, C_{b_i})$
12: **end for**
13: Calculate the 5th, median and 95th percentile for profit, costs, and optimal solution.



(a) Potential profit per source of uncertainty for one fermentation batch. Comparison between traditional operation with a fixed recipe and optimized scenario.



(b) Refinery costs separated into liquefaction and fermentation. The costs are illustrated with respect to each source of uncertainty: feed, model, and combined parameters.



(c) Solution of the optimization problem with respect to each source of uncertainty as a function of pretreatment conditions.

Figure 5: Optimal costs, potential profit, and solution of the optimization problem with respect to pretreatment temperatures, and for each source of uncertainty: feed composition, model parameters and combined.

operational point. The traditional biorefinery operation is to follow a fixed recipe with little adjustments to pretreatment conditions. This traditional recipe is most often found by offline experiments on decoupled refinery steps that do not take into account the interactions between the conversion stages and utilities prices. The results show that the traditional operation is sub-optimal from an economic point of view. In contrast, the optimization layer is capable of adapting to pretreatment temperatures and finds the optimal operation by considering the integrated process. The optimized operation is superior to a traditional recipe with a higher median profit curve at any pretreatment temperature.

At low temperatures, most of the uncertainty is due to model parameters, but it shifts after $165\,^{\circ}\mathrm{C}$ when feed uncertainty becomes dominant. The traditional operation is highly affected by feed uncertainty, while the optimized operation has a reduced uncertainty on the profit curve.

Another important result is that the optimized profit curve is flatter than the traditional curve allowing a wider range of operation with little impact on profit value. The optimal operational point can be picked as the maximum point on the median profit curve, and lies between $171\,^{\circ}\mathrm{C}$ to $176\,^{\circ}\mathrm{C}$. The optimal refinery operates at around $18\,\%$ higher profit than a traditional plant without an optimization layer.

- Figure 5b shows the refinery costs split into liquefaction and fermentation as a function of pretreatment conditions. From left to right, the uncertainty analysis is carried with respect to separate feed and model parameters (left and center plots), and combined parameters (right plot). The pretreatment costs are only due to the steam used in the thermal reactor. The biorefinery is considered to be integrated with a local power plant, possibly owned by the same company following the IBUS principle [4]. Such a design lowers the cost of steam significantly and can be neglected. A higher pretreatment temperature demands more steam but the overall increase in cost for modifying the temperature from $160\,^{\circ}\mathrm{C}$ to $180\,^{\circ}\mathrm{C}$ is negligible

compared to enzymatic hydrolysis or fermentation.

Liquefaction costs are high at low temperature because: (1) the biomatrix is not sufficiently opened to expose the whole cellulose, and (2) there is a large amount of unhydrolyzed hemicellulose, which leads to a high production of xylooligomers and xylose that inhibit the enzymatic hydrolysis further. In order to compensate for these negative effects, both the enzyme and yeast dosage are increased. The liquefaction costs decrease as the pretreatment temperature increases, which makes sense as the biomatrix opens significantly to expose cellulose, and also hemicellulose is partially removed from the enzymatic hydrolysis process.

Fermentation costs have the shape of a convex curve due to: (1) at low pretreatment temperatures a higher yeast seed could contribute to a faster digestion of sugars, which enhances the saccharification process from fermentation by reducing the C5 sugars inhibition leading to a higher ethanol yield; (2) at high temperatures the amount of inhibitors negatively affect fermentation but more yeast could compensate for the inhibitory effects of the pretreatment degradation products.

Feed uncertainty is rather constant through the entire temperature range. Uncertainty due to model parameters is high at low temperatures where the biomatrix opening highly affects the cost range. After $165\,^{\circ}\mathrm{C}$ the model uncertainty is significantly reduced becoming lower than the feed. The combined model and feed uncertainty indicate high uncertainty at low temperatures when the pretreatment is insufficient.

- Figure 5c illustrates the optimal solution as a function of pretreatment temperatures. Enzyme dosage is expressed in $\mathrm{kg/h}$, while yeast seed is given in kg. Uncertainty is higher at lower temperatures when the biomass is not sufficiently pretreated, and remains relatively constant once the biomatrix opens. Also, feed uncertainty has a higher impact than model parameters after $165\,^{\circ}\mathrm{C}$.

  Increasing pretreatment temperatures is beneficial for enzymatic hydrolysis

as it lowers the necessary enzyme dosage, but is negative for fermentation as the amount of inhibitors rises with temperature. Also, around the optimal operational point, uncertainty due to feed parameters dominates that of model parameters.

In reality there are several factors that can degrade the performance of the optimization layer, and should be accounted for in real implementation. The feed rate in this study case was set to a low value, which does not allow inhibitors accumulation in the fermentation tank. However, at higher feed rates inhibitors accumulation becomes a bottleneck, which can be counteracted by calculating an optimal feed rate profile [31]. This study also disregards the temperature dependence of the yeast performance. In reality the enzymatic hydrolysis and fermentation processes run at different optimal temperatures. The solution is to calculate a temperature profile for finding the best trade-off between the two processes [32, 33].

3.5. Global Sensitivity Analysis at the Optimal Point

The β coefficients are displayed in Table 5. The profit curve is mostly sensitive to feedstock composition, i.e. cellulosic fiber and acetyls content C_{C_S} and C_{AC_S}. $88\,\%$ of profit curve variability is explained by cellulose content changes, while $31\,\%$ is due to acetyls. A positive β coefficient indicates that an increase in cellulose content would determine an increase in profit, while a negative value decreases the profit. Acetyls have a negative impact on profit because they are responsible for acid release in pretreatment and enzymatic hydrolysis, which inhibit both the liquefaction and fermentation processes. Yield performance to convert glucose Y_{Eth_G} follows with $27\,\%$. It is also noted that unlike the local sensitivity analysis results presented above, only about 6 parameters are found to be globally influential on the profit curve (considering a threshold value for $\beta^2 >= 1\,\%$). This means that the profit curve is primarily influenced by feedstock composition.

Measuring the inflow composition accurately would help to plan the production process in advance and reduce the uncertainty on the profit curve.

Table 5: Sensitivity measure at the optimal point z_o. The linear model determination coefficient $R^2 = 0.99$.

	θ	Description	β
1	C_{C_S}	Cellulose concentration in feedstock	0.88
2	C_{AC_S}	Acetyls concentration in feedstock	-0.31
3	Y_{Eth_G}	Ethanol production from glucose	0.27
4	$q_{Max_{Ac}}$	Maximum acetate uptake rate	0.14
5	Y_{Cell_G}	Biomass growth on glucose	0.12
6	Y_{Eth_X}	Ethanol production from xylose	0.11
7	P_{MP_X}	Ethanol inhibition on xylose uptake	0.07
8	E_H	5-HMF activation energy	0.06
9	R_B	Severity factor dependency	-0.06
10	K_2	Cellulose to glucose reaction constant	0.06

3.6. Stochastic Optimization Solution

Finding the optimal point by running the process through a wide range of pretreatment temperatures requires a long computational time, and is not feasible in an industrial application. A better way is to embed the feed and model parameters uncertainty into the objective function, and pick the mean cost value:

$$
\begin{aligned}
\max_{T_{tr},F_e,M_y} \quad & \frac{1}{N}\sum_{1}^{N}\left[M_{Eth}(t_f)P_{Eth} - (F_eP_e + M_yP_y)\right] \\
\text{subject to} \quad & 0 = f(x(t),u(t)) \\
& \dot{x}_f = h(x_f,u_f) \\
150 \quad \leq \quad & T_{tr} \quad \leq \quad 210\,^{\circ}\mathrm{C} \\
10 \quad \leq \quad & F_e \quad \leq \quad 1000\,\mathrm{kg/h} \\
10 \quad \leq \quad & M_y \quad \leq \quad 1000\,\mathrm{kg}
\end{aligned}
\tag{19}
$$

This stochastic optimization is done using $N = 200$ as the number of random samples generated through LHS with correlation control. For each sample a simulation is run and the profit is assessed. After performing all N simulations the mean value of the profit is calculated, which becomes the cost function. The

following optimal solution and cost value are found:

$$z_{so} = \begin{bmatrix} T_{tr} \\ F_e \\ M_y \end{bmatrix} = \begin{bmatrix} 171.5\,^{\circ}\mathrm{C} \\ 113\,\mathrm{kg/h} \\ 147\,\mathrm{kg} \end{bmatrix} \qquad c_{so} = 7.6020 \times 10^4\,\mathrm{unitprofit} \tag{20}$$

where z_{so} is the optimal solution in the stochastic optimization case, and c_{so} is the value of the cost function. Solution (20) is relatively close to the one found in the deterministic case from (15). A slightly lower profit value is found due to the nonlinear nature of the process. The model and feed parameters are assumed to be normally distributed but when propagating through a nonlinear process the outputs change their distribution type including the mean value.

3.7. Deterministic Simulations at the Optimal Point

As a final result, a deterministic simulation is run corresponding to the optimal operational point from solution (15). The pretreatment and the enzymatic hydrolysis are continuous processes and the steady state values at the optimal point are shown in Table 6. The pretreated fibers are rich in cellulose and have a dry matter content of about $35\,\%$ as suggested by [34] for an efficient liquefaction process. Most solubles were separated from the fibers in the pretreatment process before liquefaction. The remaining hemicellulose continues to be degraded to sugars in the enzymatic hydrolysis tanks. When the level of C5 sugars increase, they strongly inhibit glucose production and a part of cellulosic fibers remain in solid state. This is why the liquefied fibers still contain cellulose before fermentation, i.e. $50\,\mathrm{g/kg}$, approximately $30\,\%$ of the initial cellulose content. The remaining cellulose continues conversion to glucose in the fermentation tank where enzymes are still active.

Figure 6 shows the fermentation batch process at the optimal point. The top plot illustrates C6 and C5 sugars depletion, ethanol production, and biomass growth. The bottom plot displays remaining cellulose and xylan conversion during simultaneous saccharification and fermentation. In the inoculum phase (first $10\,\mathrm{h}$) the yeast concentration is high but as the fed-batch phase starts,

Table 6: Steady states for pretreatment and enzymatic hydrolysis at the optimal point.

	Pretreated fibers	C5 liquid	Liquefied fibers	Fermentation
Flow/Mass	2316 kg/h	628 kg/h	2487 kg/h	220 t
Composition	1000 g/kg	1000 g/kg	1000 g/kg	1000 g/kg
Cellulose	146	1.2	50	4.4
Xylan	60	0.5	1	0
Arabinan	0	0	0	0
Lignin	85	0.7	78	60
Acetyls	16	0.1	0.1	0
Ash	6	18	5.7	7.8
Acids	1.5	4.1	16	0
Glucose	3.5	10	98	0
Xylooligomers	0.5	1.2	5.8	0.1
Xylose	10	29.7	59	0
Arabinose	5	15.5	5	0
Furfural	0.2	0.5	0.2	0
5-HMF	0.1	0.3	0.1	0
Base	0	0	6.6	9.5
Enzymes	0	0	4.9	2.4
Biomass	0	0	0	8.4
Ethanol	0	0	0	79
CO_2	0	0	0	80
Water	645	918	643	702
Other	21.2	0.2	26.6	46.4
Temperature	50 °C	50 °C	50 °C	35 °C

biomass concentration is diluted in liquefied fibers from the enzymatic hydrolysis and C5 liquid from the pretreatment process. Ethanol production has several stages: (1) formation on glucose consumption till around $100\,\mathrm{h}$, (2) production based on xylose consumption till $170\,\mathrm{h}$, (3) as xylose is depleted, its inhibition on enzymatic hydrolysis disappears and glucose production from simultaneous saccharification and fermentation is accelerated in the last $20\,\mathrm{h}$.

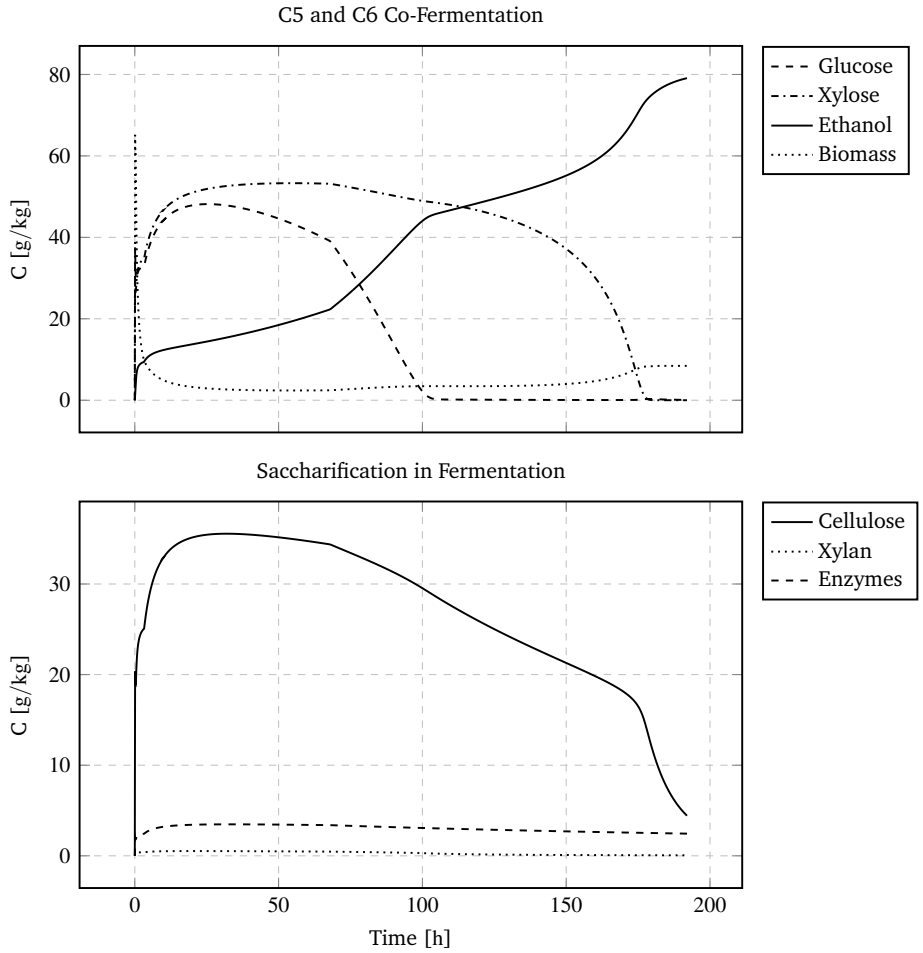


Figure 6: Simultaneous saccharification and C5-C6 co-fermentation.

4. Conclusions

This paper has presented a study on economical optimization of a large scale second generation biorefinery in a simulated environment. The proposed optimization method advices the operators of the plant at what temperature to pretreat the biomass, how many enzymes to add in the enzymatic hydrolysis tanks, and what the yeast seed should be in fermentation such that economical profit is kept high. Adjustments on these key figures need to be made on a daily basis in operation due to biomass composition and type variability, changes in consumables prices, and also due to any other disturbances that can reduce ethanol yield.

The optimization procedure is based on steady-state models (pretreatment and enzymatic hydrolysis), and dynamic fermentation model. It was found that increasing pretreatment temperature is positive for the performance of the enzymatic hydrolysis while negative for ethanol yield. Uncertainties in kinetics of pretreatment, liquefaction and fermentation were found to be negligible on the economic objective function around the optimal operational point. The main source of uncertainty was found to be the inflow feed composition.

The optimization layer reduced the uncertainty and flattened the profit curve allowing a wider range of operation with higher profit making the plant operation more robust to disturbances. The overall improvement of the optimization layer was found to be $18\,\%$ higher than the traditional plant operation.

In addition to these specific results, the paper has contributed by suggesting a generic method for biorefinery operation optimization and with quantification of sensitivity and uncertainty on earnings.

References

[1] J. Larsen, M. Ø. Haven, L. Thirup, Inbicon makes lignocellulosic ethanol a commercial reality, Biomass and Bioenergy 46 (2012) 36–45.

[2] R. C. Brown, T. Brown, S. Capareda, B. Dale, D. Edwards, V. Estes,

C. Granda, M. Holtzapple, S. Lonkar, C. E. Wyman, SBE Supplement: Lignocellulosic Biofuels, AIChE CEP (2015) 33–64.

[3] P. Cheali, J. A. Posada, K. V. Gernaey, G. Sin, Upgrading of lignocellulosic biorefinery to value-added chemicals: Sustainability and economics of bioethanol-derivatives, Biomass and Bioenergy 75 (2015) 282–300.

[4] J. Larsen, M. Ø. Petersen, L. Thirup, H. W. Li, F. K. Iversen, The IBUS Process - Lignocellulosic Bioethanol Close to a Commercial Reality, Chemical Engineering & Technology 31 (2008) 765–72.

[5] A. Sluiter, B. Hames, R. Ruiz, C. Scarlata, J. Sluiter, D. Templeton, D. Crocker, Determination of structural carbohydrates and lignin in biomass, Technical Report NREL/TP-510-42618 (2008).

[6] J. B. Kristensen, L. G. Thygesen, C. Felby, H. Jørgensen, T. Elder, Cell-wall structural changes in wheat straw pretreated for bioethanol production., Biotechnology for biofuels 1 (2008) 1–9.

[7] R. M. Prunescu, M. Blanke, G. Sin, Modelling and L1 Adaptive Control of pH in Bioethanol Enzymatic Process, in: Proceedings of the 2013 American Control Conference, Washington D.C., USA, 2013, pp. 1888–95. URL: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6580111.

[8] M. Cantarella, L. Cantarella, A. Gallifuoco, A. Spera, F. Alfani, Effect of inhibitors released during steam-explosion treatment of poplar wood on subsequent enzymatic hydrolysis and SSF, Biotechnology progress 20 (2004) 200–6.

[9] Q. Qing, B. Yang, C. E. Wyman, Xylooligomers are strong inhibitors of cellulose hydrolysis by enzymes., Bioresource technology 101 (2010) 9624–30.

[10] I. Ballesteros, M. J. Negro, J. M. Oliva, A. Cabañas, P. Manzanares, M. Ballesteros, Ethanol production from steam-explosion pretreated wheat straw., Applied biochemistry and biotechnology 129-132 (2006) 496–508.

[11] M. Ø. Petersen, J. Larsen, M. H. Thomsen, Optimization of hydrothermal pretreatment of wheat straw for production of bioethanol at low water consumption without addition of chemicals, Biomass and Bioenergy 33 (2009) 834–40.

[12] C. K. Nitsos, K. A. Matis, K. S. Triantafyllidis, Optimization of hydrothermal pretreatment of lignocellulosic biomass in the bioethanol production process, ChemSusChem 6 (2013) 110–22.

[13] E. Castro, I. U. Nieves, M. T. Mullinnix, W. J. Sagues, R. W. Hoffman, M. T. Fernández-Sandoval, Z. Tian, D. L. Rockwood, B. Tamang, L. O. Ingram, Optimization of dilute-phosphoric-acid steam pretreatment of Eucalyptus benthamii for biofuel production, Applied Energy 125 (2014) 76–83.

[14] R. Morales-Rodriguez, A. S. Meyer, K. V. Gernaey, G. Sin, A framework for model-based optimization of bioprocesses under uncertainty: Lignocellulosic ethanol production case, Computers & Chemical Engineering 42 (2012) 115–29.

[15] G. Sin, K. V. Gernaey, M. B. Neumann, M. C. M. van Loosdrecht, W. Gujer, Global sensitivity analysis in wastewater treatment plant model applications: Prioritizing sources of uncertainty, Water Research 45 (2011) 639–51.

[16] R. M. Prunescu, M. Blanke, J. M. Jensen, G. Sin, Temperature Modelling of the Biomass Pretreatment Process, in: Proceedings of the 17th Nordic Process Control Workshop, Copenhagen, Denmark, 2012, pp. 8–17. URL: http://orbit.dtu.dk/services/downloadRegister/7643840/Elektro01.pdf. doi:10.13140/2.1.1512.1287.

[17] R. M. Prunescu, M. Blanke, J. G. Jakobsen, G. Sin, Dynamic modeling and validation of a biomass hydrothermal pretreatment process - A demonstration scale study, AIChE Journal (2015).

[18] R. M. Prunescu, M. Blanke, G. Sin, Modelling and L1 Adaptive Control of Temperature in Biomass Pretreatment, in: Annual Conference on Decision and Control, Florence, Italy, 2013, pp. 3152–9. URL: http://www.nt.ntnu.no/users/skoge/prost/proceedings/cdc-2013/media/files/1221.pdf.

[19] R. M. Prunescu, G. Sin, Dynamic modeling and validation of a lignocellulosic enzymatic hydrolysis process - A demonstration scale study, Bioresource Technology 150 (2013) 393–403.

[20] C. Felby, J. Larsen, H. Jørgensen, J. Vibe-Pedersen, Enzymatic Hydrolysis of Biomasses Having High Dry Matter (DM) Content, 2006.

[21] Novozymes A/S, Cellulosic ethanol Novozymes Cellic CTec3 - Secure your plant's lowest total cost, Technical Report, 2012.

[22] C. E. Wyman, A. J. Ragauskas, Lignin Bioproducts to Enable Biofuels, Biofuels, Bioproducts and Biorefining 9 (2015) 447–9.

[23] J. J. Downs, S. Skogestad, An industrial and academic perspective on plantwide control, Annual Reviews in Control 35 (2011) 99–110.

[24] M. M. Iglesias, J. K. Huusom, K. Gernaey, State Estimation in Fermentation of Lignocellulosic Ethanol. Focus on the Use of pH Measurements, in: Proceedings of the 25th European Symposium on Computer Aided Process Engineering, 2015, pp. 1769 –74.

[25] D. E. Seborg, D. A. Mellichamp, T. F. Edgar, F. J. Doyle, Process Dynamics and Control, Process Dynamics and Control (2010) 2010.

[26] M. S. Krishnan, N. W. Ho, G. T. Tsao, Fermentation kinetics of ethanol production from glucose and xylose by recombinant Saccharomyces 1400 (pLNH33), Applied Biochemistry and Biotechnology 77-79 (1999) 373–88.

[27] R. Brun, P. Reichert, H. H. R. Künsch, Practical identifiability analysis of large environmental simulation models, Water Resources Research 37 (2001) 1015–30.

[28] G. Sin, A. S. Meyer, K. V. Gernaey, Assessing reliability of cellulose hydrolysis models to support biofuel process design - Identifiability and uncertainty analysis, Computers & Chemical Engineering 34 (2010) 1385–92.

[29] J. Helton, F. Davis, Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems, Reliability Engineering & System Safety 81 (2003) 23–69.

[30] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, S. Tarantola, Global Sensitivity Analysis. The Primer, John Wiley & Sons, Ltd, Chichester, UK, 2007. doi:10.1002/9780470725184.

[31] O. Johnsson, J. Andersson, G. Lidén, C. Johnsson, T. Hägglund, Feed rate control in fed-batch fermentations based on frequency content analysis, Biotechnology Progress 29 (2013) 817–24.

[32] S. Mutturi, G. Lidén, Effect of Temperature on Simultaneous Sacchari fi cation and Fermentation of Pretreated Spruce and Arundo, Industrial & Engineering Chemistry Research 52 (2013) 1244–51.

[33] S. Mutturi, G. Lidén, Model-based estimation of optimal temperature profile during simultaneous saccharification and fermentation of Arundo donax, Biotechnology and Bioengineering 111 (2014) 866–75.

[34] H. Jørgensen, J. Vibe-Pedersen, J. Larsen, C. Felby, Liquefaction of lignocellulose at high solids concentrations, Biotechnology and Bioengineering 96 (2007) 862–70.