# PROCESS PATHWAYS OPTIMIZATION FOR A LIGNOCELLULOSIC BIOREFINERY PRODUCING LEVULINIC ACID, SUCCINIC ACID AND ETHANOL

Aristide Giuliano<sup>a</sup>, Raffaele Cerulli<sup>b</sup>, Massimo Poletto<sup>a</sup>, Giancarlo Raiconi<sup>c</sup>, Diego Barletta<sup>a\*</sup>

<sup>a</sup>Dipartimento di Ingegneria Industriale, Università degli Studi di Salerno

Via Giovanni Paolo II, 132, 84084 Fisciano (SA), Italy <sup>b</sup>Dipartimento di Matematica, Università degli Studi di Salerno Via Giovanni Paolo II, 132, 84084 Fisciano (SA), Italy <sup>c</sup>Dipartimento di Ingegneria dell'Informazione ed Elettrica e Matematica applicata,

Università degli Studi di Salerno

Via Giovanni Paolo II, 132, 84084 Fisciano (SA), Italy

#### ABSTRACT

Process flowsheet optimization of a lignocellulosic biorefinery co-producing levulinic acid, succinic acid and ethanol was performed to maximize an economic objective function (either the Net Present Value, *NPV*, or the Internal Rate of Return, IRR, alternatively) by means of mathematical programming methods. Most promising alternative industrial processes were selected to build the superstructure of the biorefinery. A discretization method was applied to obtain a MILP approximation of the resulting MINLP master problem. *NPV* maximization for a biorefinery with hardwood feedstock provided with an optimal flowsheet with all the three products. Larger biomass allocation values resulted for levulinic acid and succinic acid (more than 40% each). A sensitivity analysis highlighted that the optimal flowsheet was significantly dependent on the economic scenario (chemical products selling price, discount rate) and on the plant scale. Finally, alternative

maximization of *NPV* and of IRR provided with different optimal flowsheets and biomass allocation to chemical products.

Keywords: biorefinery, biochemicals, process optimization, levulinic acid, succinic acid.

**DOI**: https://doi.org/10.1021/acs.iecr.6b01454

\* corresponding author Tel: +39 089 962499; Fax: +39 089 968 781; E: dbarletta@unisa.it

#### **1. Introduction**

A biorefinery is a facility or network of facilities integrating biomass conversion processes and equipment to produce transportation biofuels, power, and chemicals from biomass. This concept shows significant analogies with a conventional refinery producing multiple fuels and other products derived from oil. Progressive replacement of oil with biomass is a necessary breakthrough change to build a future biobased economy<sup>1</sup> in which fossil carbon is replaced by renewable carbon as raw material and energy source. Second generation biorefineries using lignocellulosic biomass are among the most promising options, since they have many advantages from the energetic and environmental standpoint<sup>2,3,4</sup>.

Efficient conversion of lignocellulosic materials to ethanol and value added biochemicals are still a challenging proposition. In 2004 the National Renewable Energy Laboratory listed the twelve chemical compounds more easily obtainable from lignocellulosic biomass by industrial processes<sup>5</sup>. They are value-added chemicals or precursors of other chemicals in common use. However, the commercialization of conversion technologies has been hindered by several factors including unavailability of reliable feedstock supply systems and non optimized conversion systems<sup>6</sup>. For the latter issue, the optimization by process synthesis and integration methods can help to identify the most promising pathways and to increase the profitability of biofuels and biochemicals<sup>7</sup>.

In general, the process systems design problem of biorefineries can be stated as follows: a systematic methodology is applied to choose reaction routes and corresponding process flowsheet to achieve a set of desired value-added products aiming at minimum cost (or maximum profit) and at reduced environmental impact<sup>8</sup>.

The majority of existing biorefinery processes deals with the production of biofuels and bioethanol based on thermochemical and biochemical technologies<sup>9</sup>. However, only few studies reported in the literature addressed the optimization of the biomass distribution among alternative pathways of a biorefinery co-producing alcohols (ethanol or butanol), high-value chemicals (succinic acid) and

electricity<sup>10</sup>, chemicals, like acetic acid<sup>11</sup>, beta-lactam<sup>12</sup>, i-butene<sup>7</sup> and 3-butadiene<sup>13</sup>. Significant improvements in the techno-economic feasibility can be also obtained by thermal conversion of the biomass components that cannot be easily converted by chemical or enzymatic processes<sup>14</sup>. Luo et al.<sup>11</sup> identified the best design of a multi-product lignocellulosic biorefinery producing ethanol, succinic acid, acetic acid and electricity.

The use of rigorous methods for the process unit modelling introduces significant non linearity in the constraints of superstructure optimization problems resulting in a Mixed Integer Non Linear Problem (MINLP). MINLP solution algorithms based on generalized Bender decomposition<sup>15</sup>, outer approximation<sup>16</sup>, extended cutting-plane methods<sup>17</sup> and branch-and-bound<sup>18</sup> have been developed and improved over the last decades, and implemented in generic multi-purpose solvers (e.g. MINOS, DICOPT, BARON) integrated with commercial optimization software packages such as GAMS and AMPL<sup>19</sup>. Several authors applied different methods for MINLP process optimization of biorefineries like disjunctive programming<sup>20</sup>, DICOPT program<sup>10</sup>, LINGO program<sup>21</sup>.

In some cases optimization methods include iterative methods, interfacing the optimization software with technical computing software (e.g. MATLAB)<sup>22</sup> and process simulation software (e.g. Aspen Plus)<sup>21</sup>. In all these cases the authors identified the best process pathways among the available alternatives or the best end products to maximize a techno-economic objective function.

However, the large number of integer variables necessary for large superstructures and the need to have a reasonable computational time can make the MINLP extremely challenging, even for state-of-the-art optimization software<sup>23</sup>. Furthermore, if the problem is non-convex most of the MINLP solvers risk finding a local optimum instead of the global optimum. As a result, several approaches have been proposed to convert a MINLP into an approximated Mixed Integer Linear Problem (MILP). Such problem linearization is considerably advantageous because MILP convergence solution is guaranteed by extremely fast and effective commercially available MILP solvers (e.g., CPLEX, Gurobi, Xpress).

Problem linearization can be obtained by variable discretization methods <sup>24, 25</sup>. In particular, for each non linear process variable, a vector of possible values of the variable was considered. The original variable was set equal to the sum of the product of each of these values and a binary variable. In the problem solution only one of these binary variables was allowed to be equal to 1<sup>26</sup>. This technique was applied to process optimization of a bioethanol plant from hardwood by Scott et al.<sup>27</sup>. In particular, optimization results of an original MINLP problem were compared with those of the corresponding MILP problem obtained by discretization of products of integer and continuous variables and by relaxation of non linear terms with continuous variables only.

The present work addresses the process flowsheet optimization of a multiproduct lignocellulosic biomass refinery by aiming at economic objective functions by means of mathematical programming methods. In particular, in the studied case study a hardwood biomass feedstock can be converted to levulinic acid and succinic acid<sup>5</sup>, two of the top value added chemicals, and to ethanol. The complex superstructure, as composed by alternative process routes to the three chemicals, is described in detail. Accurate estimates of the process yields and of the process unit sizes are obtained by separately addressing the modelling of hydrolysis and fermentation reactors by rigorous models accounting for the kinetics of complete reaction networks. A discretization method is applied to use the results of the reactor design in the overall flowsheet optimization of the multiproduct biorefinery and to obtain a MILP approximation of a MINLP master problem.

Two economic objective functions, namely the Net Present Value (*NPV*) and the Internal Rate of Return (IRR), are alternatively considered. In fact, *NPV* and IRR are different profitability measures which provide different results. On the one hand, *NPV* maximization favours process solutions with the highest capital cost and cash flows which often correspond to higher productivity. On the other hand, IRR maximization yields process flowsheets with the lowest capital cost and cash flows which are generally obtained for lower productivity.

The novelty of the present work is based on the application of efficient discretization and optimization methods to an integrated biorefinery plant co-producing levulinic acid, ethanol and succinic acid which was not previously addressed in the literature.

The effect of the product selling price of succinic acid<sup>11</sup> and levulinic acid<sup>28</sup> is assessed by a dedicated sensitivity analysis. In particular, a constant price value was assumed for the whole plant life. Price fluctuations, which can significantly affect overall profitability indicators<sup>13</sup>, are not considered because this issue is beyond the scope of the present study. Sensitivity analysis on the plant size is also addressed to consider different biomass volume availability in the plant site.

#### 2 Superstructure of the biorefinery process

The most promising alternative processes at the industrial scale were selected from a thorough survey of the literature and included in a complex process flowsheet to build the so-called superstructure of a biorefinery for the co-production of succinic acid, levulinic acid and ethanol. The biorefinery superstructure built for this study consists of four sections and is reported in Figure 1:

- Pretreatment;
- Hydrolysis and fermentation;
- Separation and purification;
- Thermo-chemical conversion.

Each section is formed by alternative process pathways including several process units. The main stages of each process section are described in the following. Additional data concerning the process technology and main operating conditions assumed for the modelling of each stage are summarized in Tables 1-3.

#### 2.1 Pretreatment

Biomass pretreatment generally consists of physical, chemical or thermochemical breakdown operation to obtain the three main organic components of the lignocellulosic biomass: cellulose, hemicellulose and lignin. These components can be separated on the basis of their different solubility. Hemicellulose has the highest solubility, so it is the easiest to be separated<sup>29</sup>. Using further treatments also cellulose and lignin can be separated<sup>30</sup>. Another option includes the possibility to send also the lignin to the reaction/fermentation section and then to obtain a lignin cake from the product purification section<sup>31</sup>.

The first physical treatment considered in the superstructure is the preliminary size reduction process by a mill (P1 in Figure 1). This is a common step independent of the subsequent pretreatment operation and its performances were fixed in the model. Energy requirements were calculated as a function of the biomass flowrate and of the initial and final mean particle size according to the correlation provided by Mani et al.<sup>32</sup>.

The alternative pretreatment routes considered in this study are:

- Steam explosion (P2 in Figure 1);
- Liquid hot water (P3 in Figure 1);
- Ammonia fiber explosion, AFEX (P4 in Figure 1);
- Dilute acid (P5 in Figure 1);
- Lime (P6 in Figure 1);
- Organosolvent (P7 in Figure 1);

Each route requires different operating conditions and heat duty for an effective pretreatment<sup>33</sup>. The detailed superstructure and the list of the units of the pretreatment section are reported in the Supporting Information Figure S1 and Table S1, respectively.

The steam explosion route (P2 in Figures 1 and S1) was modelled assuming a unit using medium pressure steam at 220°C with a residence time of 3 min<sup>30</sup>. After the steam explosion unit the superstructure includes a sequence of possible units to separate in a first step hemicellulose from lignin and cellulose and in a second step lignin from cellulose. As a result, the block corresponding to the steam explosion route can provide three alternative sets of outlet streams: 1) solubilised hemicellulose, solid cellulose and lignin stream; 2) solubilised hemicellulose stream and solid

cellulose and lignin stream; 3) solubilised hemicellulose stream, solid cellulose stream, solid lignin stream.

The liquid hot water route (P3 in Figures 1 and S1) was modelled assuming the use of a soaking unit using water at 190°C with a residence time of 15 min. After the soaking unit a possible solid liquid separation unit is considered<sup>34</sup>. As a result, the block corresponding to the liquid hot water route can provide two alternative sets of outlet streams: 1) solubilised hemicellulose, solid cellulose and lignin stream; 2) solubilised hemicellulose stream and solid cellulose and lignin stream.

The AFEX route (P4 in Figures 1 and S1) consists in a unit performing biomass soaking with ammonia at 90°C. Ammonia is then separated from the solubilised hemicellulose, solid cellulose and lignin stream by a flash unit and recycled to the AFEX unit after recompression<sup>35</sup>.

The dilute acid route (P5 in Figures 1 and S1) includes a soaking unit operating at 160 °C with a  $H_2SO_4$  solution. A subsequent acid neutralization stage with Ca(OH)<sub>2</sub> is necessary. After the dilute acid treatment a possible solid liquid separation unit is considered<sup>31</sup>. As a result, the block corresponding to the dilute acid route can provide two alternative sets of outlet streams: 1) solubilised hemicellulose, solid cellulose and lignin stream; 2) solubilised hemicellulose stream and solid cellulose and lignin stream.

The lime route (P6 in Figures 1 and S1) includes a soaking unit operating at 120 °C using Ca(OH)<sub>2</sub> at rates equal to 9% w/w of dry biomass. Recovery and recycle of Ca(OH)<sub>2</sub> by addition of CO<sub>2</sub> and conversion to lime is considered<sup>36</sup>. After the lime treatment a possible solid-liquid separation unit is also considered. As a result, the block corresponding to the lime route can provide two alternative sets of outlet streams: 1) solubilised hemicellulose, solid cellulose and lignin stream; 2) solubilised hemicellulose stream and solid cellulose and lignin stream.

The organosolvent route (P7 in Figures 1 and S1) consists in a soaking unit operating at 180 °C and using an ethanol water solution 60 % w/w in ethanol. The residence time is 60 min. Following this route, a distillation column is necessary to recover ethanol. A solvent make up is provided by an aqueous ethanol stream produced in the separation and purification section of the biorefinery<sup>37</sup>.

Main operating conditions and yields to products assumed for all the pretreatment process routes are summarized in Table 1.

#### 2.2 Hydrolysis and fermentation

The detailed superstructure and the list of the units of the hydrolysis and fermentation section are reported in the Supporting Information Figure S2 and Table S2, respectively.

### 2.2.1 Acid hydrolysis and levulinic acid production

Dilute acid hydrolysis (R1-R2 in Figures 1 and S2) is used to obtain glucose and xylose from cellulose and hemicellulose, respectively, and to convert glucose to levulinic acid. The reaction network includes the formation of hydroxymethylfurfural (HMF) as an intermediate in the path from cellulose to levulinic acid and the conversion of xylose to furfural. Secondary parallel reactions can take place for cellulose, glucose and xylose to undesired decomposed compounds. The main reaction network is reported in Figure 2:

Two different acid catalysts,  $H_2SO_4$  and HCl, are considered<sup>28</sup>. Acid hydrolysis is performed in a plug flow reactor at  $150^{\circ}C^{38}$ . Reaction network kinetics relevant to  $H_2SO_4^{39}$  and  $HCl^{40}$  were taken into account.

The outlet stream from the hydrolysis reactor is rich in intermediates (glucose and xylose) or in final products (levulinic acid, formic acid and furfural) depending upon the residence time considered. Of course, longer residence times favour the presence of final products. Therefore, if the conversion in the dilute acid reactor is insufficient, a second acid reactor (named levulinic acid reactor, R3-R4 in Figures 1 and S2) is used in series to produce more levulinic acid. Alternatively, dilute levulinic acid stream from the hydrolysis reactor can bypass the second acid reactor and be sent to an acid detoxification reactor (R11 in Figure S2), before feeding to fermentations in the ethanol or to succinic acid routes.

The levulinic acid reactor outlet stream is sent to an ultrafiltration filter (R10 in Figure S2) to remove solid particles, oligomers and polymers (cellulose and hemicellulose)<sup>41</sup>. Next, the filtered reaction liquor containing by products of the sugars hydrolysis and acid catalyst is conditioned by

means of a detoxification stage (acid purification) (R11 in Figure S2) before sending the sugars to fermentation<sup>31</sup>. Resulting wastewater is sent to dedicated treatment.

#### 2.2.2 Enzymatic hydrolysis

Enzymatic hydrolysis (R5 in Figures 1 and S2) is used to produce glucose and xylose from cellulose and hemicellulose using a continuous stirred-tank reactor by means of specific enzymes<sup>42</sup>. The superstructure includes two possible reactors. The first reactor can process either a mixed stream of cellulose and hemicellulose or a stream of cellulose only. Enzymatic kinetics for this case was derived from Kadam et al.<sup>43</sup>. The second reactor (R7 in Figures 1 and S2) can process a stream of hemicellulose only. The enzymatic kinetics for this case was taken from Flores-Sanchez et al.<sup>44</sup>. Two alternatives are considered to feed the enzymes to the enzymatic hydrolysis section. In one case it is assumed that the enzymes were bought from an external producer. In the second case, the assumption of an on site process section dedicated to produce cellulase enzymes<sup>31</sup> is considered. In this case part of the sugar-rich stream from the enzymatic reactor is recycled back to the enzyme production section. The two process alternatives are depicted in the flowsheet reported in Figure S3.

# 2.2.3 Hemicellulose dilute acid hydrolysis

Xylose can be also produced from hemicellulose by a dilute acid hydrolysis plug flow reactor (R6 in Figures 1 and S2) with  $H_2SO_4$  as catalyst<sup>39</sup>. The main reaction network is reported in Figure 3. After the hydrolysis, solids are separated by a filter (R61 in Figure S2) and acid agents in the stream are neutralized in a detoxification stage (R62 in Figure S2) to obtain higher pH as required for fermentations. The xylose-rich stream can be either sent to ethanol fermentation, or sent to succinic acid fermentation or split between the two units.

#### 2.2.4 Ethanol and succinic acid fermentation

The streams of dissolved sugars produced by cellulose and hemicellulose by any of the hydrolysis routes are sent to fermentation stages.

Ethanol production (R8 in Figures 1 and S2) takes place into batch fermenters in which sugars are converted into ethanol and CO<sub>2</sub> by specific microorganisms. *Zymomonas mobilis* strain ZM4

(pZB5) turns out to be the best microorganism for the co-fermentation of C6 sugars and C5 sugars<sup>45</sup>. Ethanol production reactions are summarized in Figure 4:

The ethanol fermenter is composed by a batch reactors train in order to have a continuous production despite the required long residence time<sup>31</sup>.

Succinic acid production is performed in a fermentation reactor (R9 Figures 1 and S2) fed with sugars solution and  $CO_2$  provided from the ethanol fermenter. The reaction network considered<sup>11</sup> is shown in Figure 5:

These reactions require the microorganism *Mannheimia succiniciproducens* (MBEL55E). The chosen reaction kinetics are those reported by Song et al.<sup>46</sup>.

Also the succinic acid fermenters are composed by a train of batch reactors for the same reason of the ethanol fermenter.

#### 2.3 Separation and purification

### 2.3.1 Levulinic acid purification

The purification section aims at obtaining 99 %wt levulinic acid. The detailed superstructure and the list of the units of the levulinic acid separation and purification section are reported in the Supporting Information Figure S4 and Table S4, respectively.

If HCl is the catalyst used for the acid hydrolysis the section consists in a separation stage (S1 in Figure 1) of the HCl catalyst for its recovery and a subsequent separation stage (S2) of water by distillation<sup>40</sup>. In particular the volatile HCl is recovered in the vapour stream of a flash stage (F1 in Figure S4)<sup>28</sup> that is recycled back. The resulting liquid stream made of levulinic acid, furfural, formic acid, water and other decomposed compounds is separated by a distillation column (C15 in Figure S4) where the levulinic acid is recovered as a bottom product because it is the heavier compound with a boiling temperature of 246°C at atmospheric pressure.

If  $H_2SO_4$  is the catalyst used in the acid hydrolysis, two alternative purification processes are possible:

- chromatographic separation (S3 in Figure 1) of H<sub>2</sub>SO<sub>4</sub> and subsequent separation (S4 in Figure 1) of water by distillation<sup>47</sup>;
- reactive solvent extraction (S5 in Figure 1) of organic compounds with subsequent distillation (S6 in Figure 1) on the resulting streams to recover solvent and H<sub>2</sub>SO<sub>4</sub><sup>48</sup>.

In particular, the first process route is consists in a chromatographic column (CC11 in Figure S4) with an adsorbent anionic resin used to separate  $H_2SO_4$  from the aqueous solution and all the other organic compounds<sup>47</sup>. Since this kind of chromatographic column needs a regeneration cycle, four columns in parallel were considered. The resulting aqueous solution of levulinic acid and a very small quantity of  $H_2SO_4$ , formic acid and furfural is sent to a distillation column (C14 in Figure S4). The bottom product is 99% wt levulinic acid.

The second process pathway consists in a reactive solvent extraction column (CR11 in Figure S4) with pentanol solvent. Pentyl levulinate is obtained from the levulinic acid reactive extraction while the aqueous solution of H<sub>2</sub>SO<sub>4</sub> is recovered in the raffinate stream. A distillation column will recover the acid catalyst from this latter stream. The extract stream containing mainly organic species (pentyl levulinate, furfural, pentanol and decomposed compounds) is sent to a flash unit (F2 in Figure S4) removing water and then to a train of distillation columns (C11, C12, C13 in Figure S4). Pentyl levulinate is obtained from the bottom of the first column (C11 in Figure S4) and it is sent to a reactor (RL in Figure S4) for the conversion to levulinic acid<sup>1</sup>. The top stream of the first column is sent to two additional distillation columns (C12 and C13 in Figure S4) to recover pentanol<sup>49</sup>.

### 2.3.2 Ethanol purification

The detailed superstructure and the list of the units of the ethanol separation and purification section are reported in the Supporting Information Figure S5 and Table S5, respectively. The first separation unit is a flash (E1 in Figure S5) to recover the  $CO_2$  produced in the ethanol fermentation

to be sent to the succinic acid fermenter. A ultrafiltration unit (E2 in Figure S5) is used to remove both solid particles and no hydrolysed oligomers and polymers from cellulose and hemicellulose<sup>41</sup>. Ethanol purification up to 99 % wt can be carried out by either conventional unit operations, such as distillation and extractive distillation, or innovative technologies, such as pervaporation membranes and molecular sieves.

The first purification step up to 77 % wt of ethanol can be performed by either a beer column (S7 in Figure 1 and C21 in in Figure S5)<sup>42</sup> or a pervaporation membrane by ethanol permeation (S8 in Figure 1 and E3 in Figure S5)<sup>48</sup>. A rectifying column (S9 in Figure 1 and C22 in Figure S5) achieves ethanol azeotropic composition (92.5 % wt)<sup>42</sup>.

Final purification step can be carried out by three alternative processes:

- extractive distillation (S10 in Figure 1),
- molecular sieve unit (S11 in Figure 1),
- pervaporation membranes (S12 in Figure 1).

In particular, extractive distillation requires two columns (C23 and C24 in Figure S5) and uses ethylene glycol as entrainer. The first column (C23 in Figure S5) provides pure ethanol as distillate and a mixture as bottom product, that is sent to the second column (C24 in Figure S5), where the ethylene glycol is recovered as bottom product and water as distillate<sup>50</sup>.

Molecular sieves (E5 in Figure S5) consist in zeolites adsorption units able to purify ethanol up to 99%wt<sup>51</sup>. Since molecular sieves need regeneration, two units in parallel were considered to alternate adsorption and regeneration.

Pervaporation membranes (E4 in Figure S5) are used to dehydrate the azeotropic water-ethanol mixture by water permeation<sup>52</sup>. The main process parameter is the downstream pressure of the permeate phase. A downstream pressure of 0.133 kPa was considered.

## 2.3.3 Succinic acid purification

The purification of succinic acid can be carried out alternatively by:

• reactive crystallization (S13 in Figure 1) and purification (S14 in Figure 1) with methanol<sup>53</sup>;

• solvent extraction (S15 in Figure 1) combined with distillation (S16 in Figure 1) and a crystallization<sup>54</sup>.

The detailed superstructure and the list of the units of the succinic acid separation and purification section are reported in the Supporting Information Figure S6 and Table S6, respectively. The first purification process requires a reactor (SA1 in Figure S6) to convert succinic acid to diammonium succinate by means of the reaction *a* reported in Figure 6. A crystallizing unit (CY31 in Figure S6) is used to obtain succinic acid as a solid precipitate product by the reaction *b* reported in Figure 6. Methanol solvent is used in order to purify the liquor with succinic acid crystals (SA3 in Figure S6). In fact, all residual organic compounds are dissolved in the methanol-phase and pure succinic acid is obtained. Residual succinic acid in the organic phase is also recovered by further crystallization (CY32 in Figure S6). Methanol is recovered by distillation (C31 in Figure S6). A thermal cracking (SA4 in Figure S6) allows the recovery and recycle of ammonia and ammonium sulfate<sup>53</sup>.

The alternative purification process of succinic acid consists of three units<sup>54</sup>:

- a reactive extraction column (CR31 in Figure S6) with octanol solvent to remove by-product acids from the fermentation broth;

- a vacuum distillation column (C32 in Figure S6) to eliminate residual volatile carboxylic acids such as acetic, formic and lactic acids. This operation is used to facilitate the subsequent crystallization;

- a crystallizer (CY33 in Figure S6) to obtain 99% wt pure succinic acid crystals<sup>55</sup>.

## 2.4 Thermochemical conversion

In the lignin section, the thermochemical conversion of the lignin-rich stream allows to produce high pressure steam and electricity. The lignin stream, coming either from the pretreatment section or from the separation section, is sent to two alternative process pathways:

- lignin gasification (T1 in Figure 1) coupled with a combined cycle;
- lignin combustion (T2 in Figure 1) coupled with a Rankine cycle.

The detailed superstructure and the list of the units of the thermochemical conversion section are reported in the Supporting Information Figure S7 and Table S7, respectively. In the first process route, an Integrated Gasification Combined Cycle is considered. In particular, the raw syngas from the gasification reactor (L4 in Figure S7) is sent to a gas cleaning section (L5 in Figure S7) and then to a gas turbine (GT1 in Figure S7)<sup>56</sup>. Both high pressure steam produced by heat recovery from the gasifier and flue gas at the exit of the gas turbine are sent to the heat recovery steam generator (HRSG in Figure S7). Resulting steam is sent to three steam turbines (ST4, ST5 and ST6 in Figure S7) to produce electricity and intermediate and low pressure steam<sup>57</sup>.

In the case of lignin combustion high pressure steam is produced in the boiler (L2 in Figure S7) and is sent to steam turbine train at three pressure levels (ST1, ST2 and ST3 in Figure S7)<sup>31</sup>.

#### 2.5 Potential mass and heat integration of flowsheet

The biorefinery superstructure described above offers potential mass and heat integration opportunities between different pathways. In particular, intermediate by-products could be used as feedstocks for other process units for the alternative main products. Main mass integration options are represented by:

- the CO<sub>2</sub> produced by the ethanol fermentation (R8 in Figure S2) can be used as a feed stream for the succinic acid fermentation (R9 in Figure S2);
- the sugar by-products of the acid hydrolysis (R1-R2 in Figure S2) to levulinic acid can be used as feed for the fermentation units (R8-R9 in Figure S2);
- steam produced from thermo-chemical processes (T1-T2 in Figure 1) can be used for steam explosion pretreatment (P2 in Figure 1);
- ethanol from the fermentation process and purification (Figure S5) can be used as a solvent for organosolvent pretreatment (P7 in Figure 1);
- organic residues from fermentations (R8 and R9 in Figure 1) and thermochemical section can be used into anaerobic digestion as substrates. The produced biogas can be used into the lignin

gasifier (L4 in Figure S7) or combustor (L2 in Figure S7) mixed with air for air preheating by combustion<sup>31</sup>.

Heat integration is required for the economic feasibility of a lignocellulosic biorefinery<sup>58</sup>. In fact, energy-intensive processes, such as distillation columns and pretreatment processes, can be integrated with steam generation from thermochemical processes of solid residues (lignin-rich stream). Therefore, heat integration is also addressed in the present work as detailed in section 4.3.

# 2.6 Biomass feedstock

A hardwood biomass was chosen as a representative lignocellulosic feedstock. This feedstock is a polysaccharide-rich material, with 66% wt of carbohydrate content<sup>59</sup>. Biomass composition is reported in Table 4. The cellulose weight fraction is about 50% wt, the hemicellulose weight fraction is about 16% wt, and the lignin one is 28% wt. Other compounds (mainly oils, proteins, ashes) which cannot be exploited in the proposed biorefinery superstructure are present at 6% wt.

#### 3. Superstructure mathematical modelling

From the standpoint of the mathematical description, the optimization problem of the process superstructure described in the previous section consists in:

- mass and energy balance equations;
- inherent constraints relevant to process conditions;
- design equations for the process unit sizing;
- heat integration equations;
- capital cost and operating cost equations;
- economic objective function.

In particular, the main process inherent constraints concern the exploitation of mass and energy integration opportunities described in section 2.5 to minimise the use of additional reactants and

energy utilities. For example, the  $CO_2$  needed for the succinic acid fermentation is provided by the by-product stream of the ethanol fermentation only.

#### 3.1 Material balance constraints

Several kinds of material balance equations are considered in the superstructure model.

The set of balance equations on splitters balances is given by  $^{60}$ :

$$F_{i}^{k} = \sum_{j=1}^{n_{j}} F_{i,j}^{k,IN} \mu_{j}^{k} \quad \text{for } k = 1, ..., n_{k} - 1; i = 1, ..., n_{comp}$$

$$F_{i}^{IN} = \sum_{j=1}^{n_{j}} F_{i,j}^{k,IN} \quad \text{for } k = 1, ..., n_{k} - 1; i = 1, ..., n_{comp}$$

$$F_{i}^{IN} = \sum_{k=1}^{n_{k}} F_{i}^{k} \quad \text{for } i = 1, ..., n_{comp}$$

$$\sum_{j=1}^{n_{j}} y_{j}^{k} \leq 1 \quad \text{for } k = 1, ..., n_{k} - 1$$

$$F_{i,j}^{k,IN} \leq y_{j}^{k} \text{ UB} \quad \text{for } j = 1, ..., n_{j}; \ k = 1, ..., n_{k} - 1; i = 1, ..., n_{comp}$$

$$y_{j}^{k} \in \{0, 1\}$$

$$(1)$$

where  $F_i^k$  is the component *i* mass flowrate of the outlet stream *k*;  $F_i^{IN}$  is the component *i* mass flowrate of the inlet stream;  $\mu_j^k$  are the  $n_j$  discrete values of the split fraction for the stream *k*;  $F_{i,j}^{k,IN}$ are the  $n_j$  inlet flowrate variables deriving from the disaggregation of the inlet flowrate  $F_i^{IN}$ ,  $n_k$  is the number of outlet streams,  $y_j^k$  are binary variables for the selection of outlet stream flowrates, *UB* is the upper bound for the flowrates<sup>60</sup>. When the splitter can feed more than one outlet stream, corresponding to the case in which several parallel pathways are possible, the  $\mu_j^k$  discrete values are real numbers between 0 and 1. Differently, when the splitter can feed only one outlet stream, corresponding to only one of the process pathways, the  $\mu_j^k$  discrete values reduce to the two integer values 0 and 1. This material balance formulation yield linear constraints for the continuous variables and the binary

The set of balance equations of mixer are:

$$F_i^{OUT} = \sum_{k=1}^{n_k} F_i^k$$
 for  $i = 1, ..., n_{comp}$  (2)

where  $F_i^k$  is the component i mass flowrate of the inlet stream k;  $F_i^{OUT}$  is the component *i* mass flowrate of the outlet stream.

For reactors in which the reactant *j* is converted to product *i* we have:

$$F_{i}^{OUT} = F_{j}^{IN} \eta_{i,j}(T, P, c, pH) + F_{i}^{IN} \qquad \text{for } i = 1, ..., n_{comp}$$
(3)

where  $F_i^{OUT}$  is the product *i* outlet flowrate,  $F_j^{IN}$  is the reactant *j* inlet flowrate,  $F_i^{IN}$  is the product *i* inlet flowrate,  $\eta_{i,j}(T, P, c_j, pH)$  is the yield to product *j* of reactant *i*. The yield is a function of temperature *T*, pressure *P*, catalysts concentration c, potential of hydrogen pH.

A significant improvement of the mathematical model of the superstructure regards the inclusion of the results of the rigorous reactor modelling described in section 3 in the mass balance equations of reactors:

$$F_p^{OUT} = F_r^{IN} \eta_p(\tau) + F_p^{IN}$$
(4)

where  $F_p^{OUT}$  is the product *p* outlet flowrate,  $F_r^{IN}$  is the reactant *r* inlet flowrate,  $F_p^{IN}$  is the product *p* inlet flowrate,  $\eta_p(\tau)$  is the product *p* yield obtained by equation (15). In this formulation the mass balance equation (4) is non-linear, due to the term given by the product of  $F_r^{IN}$  and of  $\eta_p(\tau)$ , in which this latter is also a non linear function of  $\tau$ . Moreover, also the reactor design equations (16) and (17), are non linear due to the presence of the product of *F* and  $\tau$ .

For separation units we have:

$$F_i^k = F_i^{IN} \zeta_{i,k}(T, P) \qquad \text{for } i = 1, \dots, n_{comp}$$
(5)

where  $F_i^k$  is the component *i* mass flowrate of the outlet stream *k*,  $F_i^N$  is the component i inlet flowrate,  $\zeta_{i,k}(T, P)$  is the fractional recovery of component *i* in the outlet stream k. This recovery is a function of temperature T and pressure P.

The following constraint was used to ensure that the  $CO_2$  produced by ethanol fermentation is sufficient as reactant for the succinic acid fermentation, as mentioned in section 2.5:

$$F_{CO2}^{EtOH, fer} \ge F_{CO2}^{SA, fer} \tag{6}$$

where  $F_{CO2}^{EtOH, fer}$  is the CO<sub>2</sub> outlet flow rate from ethanol fermenter,  $F_{CO2}^{SA, fer}$  is the CO<sub>2</sub>inlet flow rate to succinic acid fermenter.

#### 3.2 Energy balance constraints

Energy balances are also needed in order to estimate the steam and the thermal energy required by the process.

For each process unit the following general stationary energy balance equation applies:

$$\sum_{i=1}^{n_i} F_i^{IN} \Delta H_i^{IN} + Q_{GEN} + Q_{EXT} = \sum_{i=1}^{n_i} F_i^{OUT} \Delta H_i^{OUT} \quad \text{for } i = 1, ..., n_{comp}$$
(7)

Where  $Q_{GEN}$  is generated heat power during the process and  $Q_{EXT}$  is the external heat duty,  $\Delta H_i^{IN}$ and  $\Delta H_i^{OUT}$  are the specific enthalpies of component *i* at the inlet and at the outlet conditions, respectively.

For the reboilers and the condensers of the distillation columns, the heat duty was estimated, by means of the simplified relations proposed by Biegler et al.<sup>60</sup>. Under the hypotheses of constant molar fluxes and saturated liquid feed the rising vapour stream is equal to the sum of the liquid reflux stream, L, and the liquid distillate stream, D, therefore we have:

$$Q_{EXT} = (L+D)\sum_{i=1}^{n_i} x_{B,i}\lambda_i$$
(8)

for reboilers, and:

$$Q_{EXT} = -(L+D)\sum_{i=1}^{n_i} x_{D,i}\lambda_i$$
(9)

for condensers.

where,  $x_{B,i}$  and  $x_{D,i}$  are the molar fraction of the bottom and of the distillate respectively and  $\lambda_i$  is the latent heat of the component *i*. In eq. 7-9, *cp* and  $\lambda$  are considered constant with temperature.

For the mixer units, energy balances were not considered since it was always assumed the same temperature for all inlet streams and also the heat of mixing was neglected. In order to obtain the same temperature for all the inlet streams before mixing, heat exchangers were used to make the temperature difference vanish.

Energy balances on transformation units producing electricity, as steam or gas turbines were carried out in a simplified manner as:

$$E = \eta_{ISO} \eta_{MEC} (\Delta H^{IN} - \Delta H^{OUT}) F$$
<sup>(10)</sup>

Where *E* is the electricity generated by the turbine,  $\eta_{ISO}$  and  $\eta_{MEC}$  are isentropic efficiencies and mechanical efficiencies of turbine,  $\Delta H^{IN}$  and  $\Delta H^{OUT}$  are inlet and outlet specific enthalpies of stream F.

#### 3.3 Heat integration

The minimization objective on the use of external energy utilities in the whole biorefinery plant was addressed by using heat integration methods according to the pinch analysis. Using the heat duties  $Q_{EXT}$  from the solution of eqs 7-9 calculated on all the plant equipment units, the pinch theory was applied in order to minimize the required hot and cold external utilities<sup>61</sup>.Since the inlet and outlet stream temperatures of each heat exchangers are fixed according to the process requirements, a linear problem was obtained to derive the heat duty of external hot and cold utilities:

$$CP_{k} = \frac{Q_{EXT,k}}{T_{k}^{OUT} - T_{k}^{IN}} \qquad \text{for } k = 1, ..., n_{h}$$

$$Q_{j} = \sum_{k=1}^{n_{h}} CP_{k} \omega_{k} v_{j,k} (T_{j} - T_{j+1}) \qquad \text{for } j = 1, ..., n_{\text{int}}$$

$$\overline{Q}_{j} = \overline{Q}_{j-1} - Q_{j-1} \qquad \text{for } j = 1, ..., n_{\text{int}}$$

$$Q_{HOT} \ge -(\overline{Q}_{j-1} - Q_{j-1}) \qquad \text{for } j = 1, ..., n_{\text{int}}$$

$$Q_{COLD} = \overline{Q}_{n_{\text{int}}} + Q_{HOT}$$

$$(11)$$

where  $CP_k$  is the heat capacity flowrate of the stream through the heat exchanger  $k, T_k^{IN}$  and  $T_k^{OUT}$  are the temperatures of the inlet and outlet stream, respectively,  $Q_{EXT,k}$  is the heat duty of the exchanger *k*,  $n_h$  is the number of heat exchangers,  $n_{int}$  is the number of subintervals,  $T_j$  and  $T_{j+l}$  are the temperature bounds of the subinterval *j* belonging to the overall temperature range of the heat exchanger network,  $Q_j$  is the heat duty exchanged in  $\Delta T_j$ ,  $\overline{Q}_j$  is a heat duty auxiliary variable and it represent the heat necessary to heat all cold stream from  $T_j$  to  $T_{nint}$ ,  $Q_{HOT}$  and  $Q_{COLD}$  are the required heat duty of hot and cold utilities, respectively. Finally,  $\omega_k$  is equal to 1 for a hot stream and to -1 for a cold stream;  $v_{j,l}$  is equal to 1 or to 0 if the temperature subinterval  $[T_{j+1}, T_j]$  is included or not in the heat exchanger temperature range  $[\omega_k T_k^{OUT}, \omega_k T_k^{IN}]$ .

#### 4. Reactor modelling and equipment sizing

#### 4.1 Reactor modeling and sizing

Hydrolysis and fermentation reactors were modelled by rigorous methods accounting for relevant kinetic schemes to address a more accurate design of the reactors. Constant volume and temperature were assumed for all the reactors. In particular, plug flow reactors for the dilute acid hydrolysis and the levulinic acid reaction were modelled by the following molar balance equations:

$$\frac{dc_{i}}{d\tau} = \sum_{j=1}^{n_{r}} v_{i,j} r_{j}(c_{i}, c_{k}) \quad \text{for } i = 1, \dots, n_{comp}$$

$$I.C. \tau = 0 \quad c_{i} = c_{i0}$$
(12)

where  $\tau$  is the space time in the reactor,  $c_i$  and  $c_k$  are the molar concentration of component *i* (which the balance equation is considered on) and *k* (any other component affecting the reaction rate), respectively,  $r_j$  is the reaction rate of component *i* in reaction *j*,  $v_{i,j}$  is the stoichiometric coefficient for component *i* in reaction *j*, assuming negative value if component *i* is a reactant and positive values if the component *i* is a product, $c_{i0}$  is the initial concentration of component *i*,  $n_{comp}$  is the number of components and  $n_r$  is the number of reactions.

Batch reactors for ethanol and succinic acid fermentation were modelled by the following mass balance equations:

$$\frac{dw_{i}}{d\tau} = \sum_{j=1}^{n_{r}} \psi_{i,j} R_{j}(w_{i}, w_{k}) \quad \text{for } i = 1, \dots, n_{comp}$$

$$I.C. \tau = 0 \quad w_{i} = w_{i0}$$
(13)

where  $\tau$  is the batch time in the reactor,  $w_i$  and  $w_k$  are the mass concentration of component *i* (which the balance equation is considered on) and *k* (any other component affecting the reaction rate), respectively,  $R_j$  is the reaction rate of component *i* in reaction *j*,  $\psi_{i,j}$  is the component *i* yield in reaction *j*, assuming negative value if component *i* is a reactant and positive values if the component *i* is a product,  $w_{i0}$  is the initial mass concentration of component *i*,  $n_{comp}$  is the number of components and  $n_r$  is the number of reactions.

The resulting set of ordinary differential equations (12) and (13) was solved either analytically or numerically. In particular, when the reaction rates were linear functions of concentrations, analytical solutions were derived. Instead, when the reaction rates were non linear functions of concentrations, the set of ODEs was numerically solved by a variable order solver based on the numerical differentiation formulas, namely the *ode15s* function available in MATLAB<sup>®</sup>.

Continuous stirred tank reactors for enzymatic hydrolysis were modelled by the following mass balance equations:

$$w_{i} = w_{i0} + \tau \sum_{j=1}^{n_{r}} \psi_{i,j} R_{j}(w_{i}, w_{k}) \quad \text{for } i = 1, \dots, n_{comp}$$
(14)

In this case a set of algebraic non-linear equations was obtained. Its solution was numerically sought by an unconstrained non-linear optimization, the Nelder-Mead simplex algorithm, namely the *fminsearch* function available in MATLAB<sup>®</sup>.

The kinetic networks considered for each reactor and the references to the relevant kinetic model adopted are reported in Table 3.

Once the relevant balance equations, Equation (12), (13) or (14) are solved for the relevant reactors, concentration-time profiles are available and, thus, it is possible to derive the overall product yield

defined as the ratio between the converted mass of product p and the initial mass of reactant r, provided the normalization of stoichiometric coefficients:

$$\eta_p(\tau) = \frac{w_p(\tau) - w_{p0}}{w_{r0}} \quad \text{for } p = 1, \dots, n_{comp}$$
(15)

The volume of PFR and CSTR, V, was calculated as:

$$V = \frac{F\tau}{\rho} \tag{16}$$

where F is the stream mass flow rate and  $\rho$  is the stream average density.

Batch reactors for fermentation were sized taking into account the required batch time for reaction and the time for cleaning, filling and emptying of vessels,  $t_{cfe}$ . A train of reactor was considered to ensure a continuous product flowrate. Accordingly the volume of each batch reactor, V, was estimated as:

$$V = \frac{F(\tau + t_{cfe})}{\rho N_u}$$
(17)

The number of units,  $N_u$ , and the volume of batch units was estimated by minimizing the capital cost. The resulting number of units for fermentation was six. The results in terms of product yield as a function of residence time according to Equation (15) obtained by solving the set of equations (12) for PFR, the set of equations (13) for batch reactors and the set of equations (14) for CSTR are reported in the Supporting Information. In particular, Figure S8 reports the space-time profile of product yields for the dilute acid hydrolysis of cellulose and hemicellulose with HCl catalyst (Figure S8a) and with H<sub>2</sub>SO<sub>4</sub> catalyst (Figure S8b) in PF reactors. Figure S8c and S8d report the batch-time profile of product yields for the levulinic acid PF reactors converting glucose with HCl catalyst (Figure S8c) and with H<sub>2</sub>SO<sub>4</sub> catalyst (Figure S8d). Figure S9 reports the yield vs residence time operating conditions for the enzymatic hydrolysis in a CSTR converting cellulose and hemicellulose. Figures S10 and S11 report the batch time profile of product yields for the

fermentation of glucose and xylose, respectively. In particular, Figure S10 refers to ethanol fermentation and Figure S11 to succinic acid fermentation.

## 4.2 Process vessels and column sizing

The large majority of the other process vessels operated in continuous mode (pretreatment section units, flash units, reactors in the separation sections, extraction columns, crystallizers) were sized on the basis of the required residence time according to the following general equation:

$$V = \frac{\sum_{i=1}^{n_{comp}} F_{i,IN} t_R}{\rho}$$
(18)

where  $V_V$  is the volume of the vessel,  $t_R$  is the residence time.

Relevant residence time values were derived from the literature already mentioned in the superstructure description section.

With concern to distillation columns the shortcut method proposed by Biegler et al.<sup>60</sup> was used to estimate the number of trays,  $N_t$ :

$$N_{T} = 0.8 \max_{i} (N_{i}) + 0.2 \min_{i} (N_{i})$$
$$N_{i} = \frac{12.3}{\left[ (\alpha_{lk/hk} - 1)^{2/3} (1 - \xi_{i})^{1/6} \right]} \qquad \text{for } i = lk, hk$$
(19)

and the reflux ratio, R:

$$R = 0.8 \max_{i} (R_{i}) + 0.2 \min_{i} (R_{i})$$

$$R_{i} = \frac{1.38}{\left[ (\alpha_{lk/hk} - 1)^{0.9} (1 - \xi_{i})^{0.1} \right]} \quad \text{for } i = lk, hk$$
(20)

where  $\alpha_{lk/hk}$  is the relative volatility between light key component and heavy key component,  $\xi_{lk}$  and  $\xi_{hk}$  are the recovery of the light key component and of the heavy key component, respectively. For the case of distillation involving azeotropic mixtures (i.e. water ethanol mixture) the binary azeotropic system was simplified into two ideal systems by treating the azeotrope as a pseudocomponent according to the procedure proposed by Vogelpohl<sup>62</sup> which takes into account

that an azeotrope behaves like a pure component and a binary. This procedure, in fact, provides an almost constant value of the relative volatility over the column stages as for ideal binary mixtures. The column diameter was evaluated accounting for fluid dynamics limits according to the following equation:

$$D_c = \sqrt{\frac{4}{\pi} \frac{V'}{0.8U_{nf} \varepsilon \rho_g}}$$
(21)

where  $D_c$  is the diameter of the column, V' is the vapour flowrate in distillation column,  $U_{nf}$  is the linear flooding velocity in the column,  $\varepsilon$  is the fraction of the area available for vapour flow in the distillation columns,  $\rho_g$  density of the vapour stream in the distillation column.

For membrane sizing this general equation was considered:

$$F_{i,OUT} = A_m P_{m,i} \tag{22}$$

where  $A_m$  is the area of the membrane,  $P_{m,i}$  is the permeate flux of the main component *i*.

#### 4.3 Heat integration and sizing

Despite heat integration was applied to minimize the use of external utilities, a heat exchanger network based on the coupling of hot and cold process streams was not derived for the sake of simplicity. Heat integration was applied in order to minimize the need of external utilities. A detailed design of the corresponding heat exchanger network was not addressed since a simultaneous superstructure and HEN optimization would make necessary to approach a MINLP<sup>60</sup>. This was beyond the scope of the paper. As a result, a simplified design of the HEN was addressed. In particular, according to the Linhoff theory, the minimum number of exchanger units was assumed equal to be N-1, where N is the number of process streams and utilities. This assumption underestimates the number of units with respect to the heat integration case. The maximum number of exchanger units can be estimated as that for the Maximum Energy Recovery (MER) as  $(N_b-1)+(N_a-1)$  where N<sub>b</sub> and N<sub>a</sub> are the number of streams below the pinch and the number of streams

above the pinch respectively. In the worst case we can assume that all the streams are both below the pinch and above the pinch. This yields  $N = N_b = N_a$  and thus the maximum number of exchanger units become 2(N-1).

Heat exchanger areas were evaluated by using linear equations, assuming overall heat transfer coefficients,  $U_k$ , as a function of the type (phase, organic or inorganic species) of exchanging streams and mean temperature differences  $\Delta T$  constants (40°C) for all heat exchangers, since the stream coupling are unknown:

$$A_k = \frac{Q_{EXT,k}}{U_k \Delta T}$$
(23)

As a result of the above described assumptions, the sizes of all the pieces of equipment (including distillation columns and heat exchangers) resulted linearly proportional to the stream mass flow rate or to the heat flow rate.

#### 5. Capital cost constraints and economic analysis

Capital costs of equipment were estimated by power law of capacities of the specific unit to account for the real scale:

$$CC_{k} = CC_{k,0} \left(\frac{S_{k}}{S_{k,0}}\right)^{m_{k}}$$
(24)

where  $CC_{k,0}$  is the capital cost of the equipment item *k* with the base capacity or size  $S_{k,0}$ , and  $CC_k$  is the capital cost the equipment item *k* with the real capacity or size  $S_k$ ,  $m_k$  is the sizing exponent for the kind of unit *k*. The main relevant economic data for the equipment cost were retrieved from the work by Hamelinck et al.<sup>56</sup> on the techno-economic performance of lignocellulosic biorefineries producing bioethanol. The capital costs of the purification equipment units were calculated separately. A piecewise linearization of eq. 24 was performed in order to obtain a linear correlation between capital costs and unit capacity or size. Chemical Engineering Plant Cost Index (CEPCI) was applied to update to the year of the analysis the equipment cost estimated years ago:

$$ECC_{k} = CC_{k} \frac{CEPCI_{2014}}{CEPCI_{ref}}$$
(25)

where  $ECC_k$  is direct capital cost in the year of interest for the analysis, CEPCI<sub>2014</sub> and CEPCI<sub>ref</sub> are Chemical Engineering Plant Cost Index in 2014 and in the costing reference year, respectively.

The total project investment cost (TIC) was calculated by adding to the bare module cost ECC and the auxiliary costs. The latter were assumed as proportional to ECC by means of specific factors as follows :

$$TIC = \sum_{k \in Eqp} ECC_k (1 + K_k)$$
(26)

where  $K_k$  is a cost factor accounting for equipment purchase, installation, instrumentation and controls, piping, electrical systems, buildings including services, engineering, construction, legal and contractors fees, project contingency, land and working capital. Values of  $K_k$  were set to those suggested by Biegler et al.<sup>60</sup>.

# 5.1 Total annual cost and revenues

The total annual cost (*TAC*) of the process is the sum of operational cost (*OPC*) and feedstock purchasing cost (*FSC*):

$$TAC = OPC + FSC \tag{27}$$

The operational cost (*OPC*) includes the maintenance and labour cost (*MLC*), process utility (steam) cost (*STMC*), natural gas cost (*NGC*), and power cost (*PWC*):

$$OPC = MLC + STMC + NGC + PWC$$
<sup>(28)</sup>

The cost of maintenance and labour *MLC* is evaluated as proportional to the total investment cost (*TIC*) by a specified cost factor  $f_c$ :

$$MLC = f_c TIC \tag{29}$$

Relevant prices of utilities were obtained from Humbird et al.<sup>31</sup>.

The process revenues come from the sales of levulinic acid, ethanol, succinic acid and electricity at market price (updated to September 2014). As a result, the revenues are given by:

$$\operatorname{Rev} = \sum_{p=1}^{n_p} F_p p_p \tag{30}$$

where  $F_p$  is the mass production rate of product p,  $p_p$  is the unit market price (\$/kg) of product  $p^{63}$ ,  $n_p$  is the number of products. The market prices are given in Table 4.

#### 5.2 Profitability analysis objective functions

Given the life span of the project,  $t_{ls}$ , and made the hypothesis that the plant construction takes 3, the cash flow for the year *i* is defined as:

$$CF_i = -f_i TIC + g_i WC + (Rev - TAC)(1 - t) + DPt$$
(31)

Where  $f_i$  is the fraction of *TIC* spent during year *i*, describing the investment distribution over the years, *WC* is the working capital, *t* is the tax rate and *DP* is the depreciation,  $g_i$  is a parameter equal to -1 for i = 3 (the year before the plant start-up), 1 for  $i=t_{1s}$ ,0 for all other values of *i*. In particular, a straight-line depreciation for ten years was assumed.

Two economic objective functions based on the cash flow analysis were considered. The first was the Net Present Value defined as:

$$NPV = \sum_{i=0}^{t_{ls}} \frac{CF_i}{(1+r)^i}$$
(32)

Where  $CF_i$  is the cash flow for the year i, *r* is the annual discount rate.

#### 6. MILP discretization methodology

As a consequence, the resulting optimization mathematical model is a Mixed Integer Non Linear Problem (MINLP) with the integer variables for the selection of alternative technologies:

$$\begin{cases} \max z = c(x^{L}, x^{N}, y) \\ h(x^{L}, x^{N}, y) = 0 \\ g(x^{L}, x^{N}, y) \le 0 \\ x^{L}, x^{N} \in \mathbb{R}^{+}, y \in \{0, 1\} \end{cases}$$
(33)

where z is the economic objective function,  $x^{L}$  is the vector of continuous linear variables corresponding to the mass flowrates and heat duties,  $x^{N}$  is the vector of continuous non-linear variables representing the reactor variables, residence times and product yields, y is the vector of binary variables necessary for the logics of decisions.

The equality constraints  $h(x^L, x^N, y) = 0$  represent mass and energy balances constraints and cost related constraints. The inequality constraints  $g(x^L, x^N, y) \le 0$  are related to design specifications (i.e., capacity limits and upper and lower bounds on process variables, etc.).

As a result, the problem statement reported in Equation (31) can be rearranged as follows:

$$\begin{cases} \max z = a(x^{N})x^{L} + by \\ A(x^{N})x^{L} + By = 0 \\ C(x^{N})x^{L} + Dy \le 0 \\ x^{L}, x^{N} \in \mathbb{R}^{+}, y \in \{0, 1\} \end{cases}$$
(34)

Where  $a(x^N)$  is a vector of non linear functions of the continuous variables  $x^N$ ,  $A(x^N)$  and  $C(x^N)$  are matrices of non linear functions of the continuous variables  $x^N$ , b is a vector of coefficients, B and D are matrices of coefficients.

The search of the optimal solution of a MINLP problem might be very difficult to address. As a result, a variable discretization method was applied to linearize the problem<sup>25</sup>. In particular, for each non linear variable  $x^N$  (residence time and process yield), a vector of possible values  $\zeta_i$  of the variable was considered. The original variable was set equal to the sum of the product of each of these values and a binary variable. In the problem solution only one of these binary variables was allowed to be equal to 1. This transformation can be expressed in mathematical terms as follows:

$$\begin{aligned} x_{j}^{N} &= \sum_{i=1}^{n_{j}} \zeta_{j,i} y_{j,i}^{d} \\ a_{j}(x_{j}^{N}) &= \sum_{i=1}^{n_{j}} a_{j}(\zeta_{j,i}) y_{j,i}^{d} = \sum_{i=1}^{n_{j}} \alpha_{j,i} y_{j,i}^{d} \\ x_{k}^{L} &= \sum_{i=1}^{n_{j}} x_{k,i}^{L} \\ x_{k,i}^{L} &\leq y_{j,i}^{d} UB \\ a_{j}(x_{j}^{N}) x_{k}^{L} &= \sum_{i=1}^{n_{j}} \alpha_{j,i} x_{k,i}^{L} \\ \sum_{i=1}^{n_{j}} y_{j,i}^{d} &= 1 \\ y_{j,i}^{d} \in \{0,1\} \end{aligned}$$
(35)

where  $\zeta_{j,i}$  are the discrete values assumed for the non-linear variable  $x_j^N$ ,  $n_j$  is the number of discrete values assumed for the non linear variable  $x_j^N$ ,  $y_{ji}^d$  are the binary variables of the discretization,  $\alpha_{ij}$ . are the discrete values of the function  $a_j(x_j^N)$ ,  $x_{k,i}^L$  are additional linear variables, *UB* is the upper bound.

As an example of the application of this discretization formulation we can consider the mass balance equation on a reactor:

$$F_p^{OUT} = F_r^{IN} \eta_p(\tau) \tag{36}$$

For the sake of simplicity, this equation differs from Equation (4) only for the absence of the term related to the presence of the product component in the inlet stream. By applying the discretization of Equation (34), we have:

$$\tau = \sum_{i=1}^{n_d} \overline{\tau}_i y_i^d$$

$$\eta_p(\tau) = \sum_{i=1}^{n_d} \eta_p(\overline{\tau}_i) y_i^d = \sum_{i=1}^{n_d} \overline{\eta}_{p,i} y_i^d$$

$$F_p^{OUT} = \sum_{i=1}^{n_d} F_{r,i}^{IN} \overline{\eta}_{p,i}$$

$$F_r^{IN} = \sum_{i=1}^{n_d} F_{r,i}^{IN}$$

$$\sum_{i=1}^{n_d} y_i^d \le 1$$

$$F_{r,i}^{IN} \le y_i^d \text{ UB}$$

$$y_i^d \in \{0,1\}$$
(37)

where  $\bar{\tau}_i$  are the discrete values assumed for  $\tau$ ,  $n_d$  is the number of discrete values,  $y_i^d$  are the binary variables of the discretization,  $\bar{\eta}_{p,i}$ . are the discrete values of the function  $\eta_p(\tau)$ ,  $F_{r,i}^{IN}$  are additional linear variables. The discrete values for  $\bar{\tau}_i$  and  $\bar{\eta}_{p,i}$  were derived from the functions reported in Figures S8, S11.

This transformation allowed obtaining a mixed linear problem (MILP) as an approximation of the initial MINLP problem:

$$\begin{cases} \max z = \alpha x^{L} + \beta y + \gamma y^{d} \\ Ax^{L} + By + Ey^{d} + \delta = 0 \\ \Gamma x^{L} + \Delta y + Hy^{d} + \varepsilon \le 0 \\ x^{L} \in \mathbb{R}^{+} \ y, \ y^{d} \in \{0, 1\} \end{cases}$$

$$(38)$$

where  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ ,  $\varepsilon$  are vectors of coefficients, A, B,  $\Gamma$ ,  $\Delta$ , E, H are matrices of coefficients.

On the one hand, this simplification allowed using more efficient solution methods for MILP optimization problems. On the other hand, a significant increase of the number of binary variables and real variables was determined. In fact, for each non-linear variable  $x^N$  a set of several binary variables  $y^d$  (typically 10) was introduced by the discretization procedure.

The MINLP master problem included 2098 continuous variables and 23 binary variables and 1444 equality and inequality constraints. The resulting MILP after discretization consisted in 7106 continuous variables, 2092 integer variables and 11982 equality and inequality constraints.

The resulting MILP was solved by means of AMPL software with IBM CPLEX optimization solver based on the simplex method and "branch & bound" or "branch & cut" methods<sup>60</sup>. The latter methods seek the MILP solution by first searching the solution of the corresponding "relaxed" linear programming problem with all real variables and then forcing one by one all the binary variables to assume either the value 0 or the value 1. Optimization model results were obtained using a workstation with an Intel Xeon 2 GHz CPU/8GB RAM. Computation time was of the order of 2h.

#### 6. Results

The optimal process flowsheet was sought for different case studies involving a biorefinery fed with hardwood and by assuming a feedstock flow rate of 50 t/h and 7200 plant operating hours per year. The hardwood composition and the main economic parameters for the optimization analysis are reported in Table 4. References for model assumptions regarding each process sections are reported in Tables 1-3. Firstly, the base case with the maximization of the Net Present Value was studied. Then, a sensitivity analysis was performed on the plant size and the main economic parameters. The economic parameters used in the base case are shown in Table 4.

### 6.1 Base case: maximization of the NPV

The optimal flowsheet obtained from the superstructure by maximizing the *NPV* is reported in Figure 7a. As a result, the biorefinery includes lines for all the three possible products, which are levulinic acid, ethanol and succinic acid. However, the distribution of the biomass allocation to products (kg of biomass in a certain product line per kg of biomass feed), are reported in Table 5, and the product yields (kg of product per kg of biomass feed), are reported in Table 5. In particular, according to this complete model biomass allocation values to levulinic acid (46%) and to succinic

acid (42%) are comparable. Considering yields, levulinic acid is the main product (8.9%), followed by succinic acid (5.7%) and ethanol (2.2%). This result can be explained considering that cellulose is the main component (50%) of the biomass and the levulinic acid process route can use only the cellulose fraction. Instead, process routes to ethanol and succinic acid exploit both glucose and xylose streams derived from cellulose and hemicellulose. Ethanol minor contribution to the revenues and thus to the maximization of the *NPV* causes a much lower production than the other two chemicals. Given this consideration, inspection of the flowsheet suggests that the flow rate of the produced ethanol is determined by the flow rate of the  $CO_2$  necessary for the succinic acid fermentation. Global yield, considering all the products, is about 16.8%. This value is lower than the value of the yield to ethanol from hardwood biomass commonly found in the literature (about 24%)<sup>64</sup>. This discrepancy is due to lower yield values specific of the routes to succinic acid and levulinic acid.

In the case of the results reported in Figure 7a and Table 5 the enzymes come from a dedicated on site production section. The reason is that the local manufacturing cost turns out to be lower than the enzyme purchase cost.

Heat integration resulted in a much more significant reduction of the required utilities. The global exchanged heat in the integrated network of hot and cold process streams was about 70 MWt. This heat integration allowed avoiding about 126 t/h of LP/HP steam as hot utilities.

Figure 8 reports the distribution of the total investment cost (TIC) of the different plant sections corresponding to the optimal flowsheet reported in Figure 7. In particular, the total investment is 413.4 M\$. This result confirmed that rigorous reactor modelling taking into account the complete kinetic networks is necessary for the accurate design of these units and for the overall process optimization. The total investment cost is due to hydrolysis and fermentation for the 35%, to pretreatment for 26%, to thermochemical conversion for 20%, to separations and purification for 15%, to heat exchanger network for 4%. The highest cost of the hydrolysis and fermentation section includes also the enzyme production unit which contribute with about 30% of the section cost.

Figure 9 reports the total annual cost distribution. Inspection of the figure suggests that biomass raw materials, reactants, utilities and other manufacturing costs cover 12%, 17%, 26% and 45% of the total, respectively.

The maximized Net Present Value *NPV* obtained for the optimal flowsheet was 677 M\$. The corresponding Internal Rate of Return, *IRR*, given by the solution of the following equation:

$$\sum_{i=0}^{t_{ls}} \frac{CF_i}{(1+IRR)^i} = 0$$
(37)

turned out to be 25%.

#### 6.2 Sensitivity analysis

#### 6.2.1 Plant size

In general, the plant size and the production rate play a significant role on the profitability of an industrial process. As a result, the effect of the plant size, in terms of the biomass feedstock rate, on the optimal flowsheet, biomass allocation and product yield was studied. The base case feedstock rate (50 t/h) was decreased and increased by a factor of about 3. Biomass allocation, reported in Table 5, is the same for 50 and 150 t/h biomass feed rate. Instead, levulinic acid is not produced at all for a biomass feedstock rate of 15 t/h. As a result, a plant size threshold value should exist in the range 15-50 t/h to justify the existence of a levulinic acid production route. Consistently, the optimal flowsheet for the 15 t/h biomass feed rate (Figure 7b) does not include the process route to levulinic acid. Moreover, further inspection of the flowsheet reveals that steam explosion (P2) was selected as the most convenient pretreatment process. In fact, technologies with lower investment cost are preferred for a smaller plant size due to the effect of the economies of scale. Finally, it is worth pointing out that enzymes for the enzymatic hydrolysis (R5) are not produced in situ, but externally purchased. This difference with the flowsheet for 50 t/h biomass feed rate can be explained considering that for smaller plant size it is more convenient for the *NPV* maximization to reduce the investment cost for process units and increase the manufacturing cost.

Table 5 reports the results of the economic analysis as a function of the plant size. As expected, the plant size has a significant effect on the IRR and on the profit per mass unit of biomass. In fact, the ratio between the *NPV* and the total biomass fed to the plant during its lifetime turns out to be 72, 75 and 119 \$/t of biomass for the three plant size corresponding to biomass feed rates 15, 50 and 150 t/h. Of course, it is worth pointing out that in this analysis it was not considered at all that the biomass supply cost could vary significantly with the plant size. In fact, the larger is the amount of biomass to feed, the wider could be the geographical area of the suppliers. Thus, the mobilisation, the logistics as well as the wider and more differentiated sources could affect significantly the biomass cost depending on the scenario. This point, however, is beyond the scope of this work and deserves further studies.

### 6.2.2 Products selling price

Sensitivity analysis on the product selling price of succinic acid and levulinic acid was carried out. In particular, a constant price value was assumed for the whole plant life. Price fluctuations, which can significantly affect overall profitability indicators<sup>13</sup>, were not considered because this issue is beyond the scope of the present study. A  $\pm$ 50 % variation was considered and its effect on the optimal flowsheet, biomass allocation and product yield was assessed. Table 5 reports the biomass allocation for varying price of succinic acid and of levulinic acid, respectively. The values reported in Table 5 point out that by reducing the succinic acid price by 50%, the biomass allocation to levulinic acid increases up to about 90%. The residual 10% biomass is used for succinic acid and ethanol production. Instead, by increasing the succinic acid price by 50%, levulinic acid is not produced at all. This result is obtained also if the levulinic price decreases by 50% (Table 5). Instead, if the levulinic acid price increases by 50%, levulinic acid product. This demonstrates once more that ethanol production is strictly linked to the succinic acid production for the CO<sub>2</sub> mass integration. The optimal flowsheet corresponding to this case is reported in Figure 7c. A significant difference with respect to previous flowsheet is that all the cellulose and hemicellulose

are converted together in a H<sub>2</sub>SO<sub>4</sub> dilute acid hydrolysis (R2). The latter produces levulinic acid for separation and purification and glucose and xylose for ethanol and succinic acid fermentation.

Table 5 reports also the results of the economic analysis obtained by varying the biochemicals selling prices. On the one hand, as expected the *NPV* increases with increasing both biochemicals price. On the other hand, the IRR increases with increasing succinic acid price and with decreasing levulinic acid price. This evidence can be explained considering that by increasing by 50% the succinic acid price, the revenues increase by about 30% while the total investment cost decreases by about 40%, due to the absence of the levulinic acid process route. Differently, by increasing by 50% the levulinic acid price, the revenues increase by about 100% while the TIC increases by more than 220%, due to the increase of the size of the levulinic acid production units.

#### 6.2.3 Discount rate

A sensitivity analysis was performed also on the discount rate r which is dependent on the economic scenario. As a result, the maximization of the Net Present Value was pursued for several values of r. This analysis was important for two reasons. The first was to take into account possible changes of this economic index. The second one was to establish the dependence of objective function from this economic parameter.

The results are reported in terms of optimal *NPV* and corresponding *IRR* shown in Figure 10. As expected the *NPV* decreases with r. Differently, the *IRR* is not affected significantly by the discount rate for r values in the range 3-10%. Instead, IRR increases with increasing discount rate between 12 and 15%. This result corresponds to the evidence that the optimal flowsheet turned out to be the same of the base case (Figure 7a) for r values in the range 3-10%.Instead, for r between 12 and 15% the optimal flowsheet does not include the production of levulinic acid as in the above mentioned cases of small plant size and of high succinic acid price. The corresponding optimal flowsheet is similar to that reported in Figure 7b. The levulinic acid process units were not present in the optimal solution to reduce the capital cost for the *NPV* maximization. In fact, the negative contribution of

the TIC to the *NPV* is heavier for higher discount rate. On the whole, despite the *NPV* decreases, *IRR* increases due to the reduction of the TIC for r between 12 and 15%.

#### 6.3 Maximization of the IRR

In the previous section it was highlighted that the optimal flowsheet obtained by maximizing the *NPV* is affected by the discount rate, which is very dependent on the economic scenario and period. Differently, the internal rate of return *IRR* is not influenced by the discount rate. Of course, also this profitability indicator is affected by long term uncertainty as demonstrated by dedicated studies<sup>13</sup>. However, this issue was not addressed in this work. As a result, in this final part of the work, the *IRR*, calculated according to Equation (37), was chosen as the objective function to maximize. The *IRR* maximization problem was addressed by using the iterative algorithm proposed by Pintaric and Kravanja<sup>65</sup>.

The optimal flowsheet obtained by maximizing the *IRR* is like that already reported in Figure 7b. This result corresponds to a biorefinery producing succinic acid and ethanol and no levulinic acid. As described above, the optimal flowsheet includes steam explosion (P2) for the biomass pretreatment and enzymatic hydrolysis (R5) to obtain the sugars. Enzymes are not produced in situ, but are purchased. This result can be explained considering that the *IRR* maximization provides optimal flowsheets which minimize the investment costs<sup>66</sup>. As a result, parallel process routes which require additional process units are often cancelled and pathways allowing less pieces of equipment with a larger size are selected by the optimization algorithm. Thus, the most significant effect in the biorefinery flowsheet is the elimination of the levulinic acid production. Biomass allocation values and product yield values obtained by *NPV* maximization. Biomass allocation to succinic acid and to ethanol is 74% and 26%, respectively (Figure 11a). The global product yield is 13.9% and is lower than that for maximum *NPV* (16.8%). In particular, the succinic acid yield is

10% and the ethanol yield is 3.9%. This result can be explained by considering that higher *IRR* can be obtained with lower revenues (lower productivity) and lower *TIC*.

Figure 8 reports the *TIC* for the two economic objective functions. It is confirmed that, using the *IRR* as objective function, lower *TIC* values are obtained. This result agrees also with the general finding by Kasas et al. (2011) that the maximization of *IRR* is significantly related to the minimization of *TIC*. Conversely, the maximization of *NPV* determines a larger production plant with higher revenues and, thus, with higher *TIC*. For the *IRR* maximization case a lower total product yield value corresponds to process units with smaller size in the pretreatment, hydrolysis and fermentation, separation and purification sections. This determines lower *TIC*. In particular, *TIC* required for the lignocellulosic biorefinery by maximizing *NPV* are 413 M\$, while less than half this value (179 M\$) was obtained by maximizing *IRR*. This significant reduction corresponds to the process flowsheet:

- No acid hydrolysis section and levulinic acid purification section;
- Steam explosion (P2) as pretreatment instead of dilute acid (P5);
- No enzyme production section.

Figure 9 reports the comparison of the total annual cost, *TAC*, distribution for the two economic objective functions. Inspection of the figure suggests that the costs for enzymes and reactants are higher for the case without levulinic acid production. In fact, enzymes used in the hydrolysis of the process routes to ethanol and succinic acid are more expensive than the acid used in the levulinic acid production. Differences in other manufacturing costs values, mainly consisting in labour and maintenance costs, are related to *TIC* different values. In particular, maintenance costs were estimated as 10% of total investment cost.

Of course, different *NPV* were obtained by considering the *NPV* itself as the objective function or the *IRR* as the objective function. In both cases and 8% annual discount rate is applied to calculate *NPV*. In particular, the maximum *NPV* turned out to be 620 M\$, while the *NPV* for maximum *IRR* was 556 M\$. Consistently, *IRR* was 25% and 36% by maximizing *NPV* and *IRR*, respectively.

#### 7. Conclusions

The process optimization of a multi-product biorefinery was carried out by a model embedding the design optimization of the hydrolysis and fermentation reactors by means of rigorous kinetics modelling. This work demonstrated that the coupling of the reactor modelling results with the overall process superstructure modelling can be successfully performed by a discretization procedure. In particular, this method allowed to reduce the MINLP master problem to a MILP problem and to successfully approach it by a linear optimization solver. Despite the increase by a factor of 100 of the binary variables and by a factor of 8 of the constraints due to the adopted discretization procedure, a reasonable computational time for the search of the optimal solution was required.

The optimal flowsheet obtained from the superstructure by maximizing the *NPV* consists in a biorefinery producing all the three possible chemicals, levulinic acid, succinic acid and ethanol. Comparable biomass allocation values were obtained for levulinic acid and succinic acid (more than 40% each). As expected, ethanol has the lowest biomass allocation due to its significantly lower price. In particular, the flow rate of the produced ethanol was mainly determined by the flow rate of  $CO_2$  necessary for the succinic acid fermentation. The global yield of the biorefinery, considering all the products, (about 16.8%) was lower than the common yield to ethanol from hardwood biomass. This discrepancy is due to lower yield values to succinic acid and levulinic acid.

The results of the sensitivity analysis highlighted that the optimal flowsheet and the relevant technical and economic performances are significantly dependent on the economic scenario (chemical products selling price, discount rate) and on the plant scale. In particular, threshold values of the high value chemicals price, levulinic acid and succinic acid, can be derived for the biomass allocation distribution and process flowsheet.

Finally, as expected process optimization achieved by maximizing two different economic objective functions, *NPV* and *IRR*, provided different optimal flowsheets and biomass allocation to chemical

products. The maximization of *NPV* determined a larger production plant with higher product yields and revenues, but also higher *TIC*. *IRR* maximization provided optimal flowsheets which minimize the investment costs. For the present multiproduct biorefinery case, this resulted in the elimination of the process route to levulinic acid, the reduction of the biorefinery global product yield and productivity.

## **Supporting Information**

The paper is accompanied by supporting information. In these file all the flowsheets of the biorefinery process sections are reported in detail. Moreover, for each reactor the kinetic network for the modelling of each reactor and the relevant results in terms of product yields as a function of residence time are reported. Finally, assumed product yields for the purification sections are also reported.

# ABBREVIATIONS

| <i>CEPCI</i> <sub>2014</sub> | Chemical Engineering Plant Cost Index in 2014               |
|------------------------------|---|
| <b>CEPCI</b> <sub>ref</sub>  | Chemical Engineering Plant Cost Index in the reference year |
| CF                           | Cash Flow, \$/y   |
| DP                           | depreciation, %   |
| FSC                          | feedstock purchasing cost, \$/y                             |
| hk                           | heavy key component   |
| IRR                          | Internal Rate of Return, %                                  |
| lk                           | light key component   |
| MILP                         | Mixed Integer Linear Problem                                |
| MINLP                        | Mixed Integer Non Linear Problem                            |
| MLC                          | maintenance and labour cost, \$/y                           |
| NGC                          | natural gas cost, \$/y                                      |
| NPV                          | Net Present Value, \$                                       |
| OPC                          | operational cost, \$/y                                      |
| р                            | product -   |
| Р                            | pressure, bar   |
| PBP                          | Payback period, y   |
| pН                           | potential of hydrogen                                       |
| PWC                          | power cost, \$/y  |
| Rev                          | Process revenues, \$/y                                      |
| ROI                          | Return on investment, %                                     |
| STMC                         | process utility cost, \$/y                                  |
| TAC                          | Total Annual Cost, \$/y                                     |
| TDC                          | total depreciable capital, \$                               |
|                              |   |

| tment Cost, \$ |
|----------------|
|                |

WC Working Capital, \$

# SYMBOLS

| Continuous   | variables $A_k$ required area for the heat exchanger k, $m^2$                                     |
|--|---|
| $A_m$  | required area for the membranes, $m^2$  |
| $CC_k$   | Capital Cost of the unit k, \$  |
| $CP_k$   | heat capacity flowrate of the stream through the heat exchanger $k$ , $J/K$                       |
| D  | liquid distillate stream, kg/h  |
| $D_c$  | diameter of the distillation columns, m   |
| $ECC_k$  | direct capital cost, \$   |
| F  | stream mass flow rate, t/h  |
| $F_{CO2}^{EtOH,fer}$   | $CO_2$ outlet flow rate from ethanol fermenter, kg/h  |
| $F_{\scriptscriptstyle CO2}^{\scriptscriptstyle S\!A,f\!er}$ | CO <sub>2</sub> inlet flow rate to succinic acid fermenter, kg/h                                  |
| $F_i^{OUT}$  | component i mass flowrate of the outlet stream k, kg/h  |
| $F_i^{I\!N}$   | component i mass flowrate of the inlet stream, kg/h   |
| $F_{i,j}^{k,I\!N}$   | inlet flowrate variables deriving from the disaggregation of the inlet flowrate $F_i^{IN}$ , kg/h |
| $F_p$  | mass production rate of product p, kg/h   |
| $F_p^{IN}$   | product p inlet flowrate  |
| $F_p^{OUT}$  | product p outlet flowrate   |
| $F_r^{IN}$   | reactant r inlet flowrate   |
| $F_{r,i}^{IN}$   | additional linear variables   |
| L  | liquid reflux stream to column, kg/h  |
| $N_T$  | number of trays of a distillation column  |

| R                               | reflux ratio of a distillation column                             |
|---------------------------------|---|
| $Q_j$                           | heat duty exchanged, W  |
| $\overline{\mathcal{Q}}_{_{j}}$ | heat duty auxiliary variable, W                                   |
| $Q_{COLD}$                      | required heat duty of cold utilities, W                           |
| $Q_{EXT}$                       | external heat duty, W   |
| $Q_{_{EXT,k}}$                  | heat duty of the exchanger k, W                                   |
| $Q_{GEN}$                       | generated heat power, W   |
| <i>Q</i> нот                    | required heat duty of hot utilities, W                            |
| $S_k$                           | Size of the unit k  |
| V                               | volume of PFR and CSTR, $m^3$                                     |
| V'                              | vapour flowrate in distillation columns, kg/h                     |
| $V_V$                           | volume for the vessels, $m^3$                                     |
| x                               | Continuous variables  |
| $x_{k,i}^L$                     | additional linear variables                                       |
| $x^L$                           | vector of continuous linear variables                             |
| $x^N$                           | Non linear continuous variables                                   |
| $x^N$                           | vector of continuous non linear variables                         |
| <b>r</b> j                      | reaction rate of component i in reaction j                        |
| $R_j$                           | reaction rate of component i in reaction j                        |
| Z                               | economic objective function                                       |
| $a(x^N)$                        | vector of non linear functions of the continuous variables $x^N$  |
| $\Delta H_i^{IN}$               | specific enthalpies of component i at the inlet conditions, J/kg  |
| $\Delta H_i^{OUT}$              | specific enthalpies of component i at the outlet conditions, J/kg |
| $\eta_{i,j}$                    | yield to product j of reactant i                                  |
| ${\pmb \eta}_p$                 | overall product yield of product p                                |

# Parameters

| Α                       | vector matrices of non linear functions of the continuous variables $x^N$ |
|-------------------------|---|
| b                       | vector of coefficients  |
| В                       | matrix of coefficients  |
| С                       | catalysts concentration, M  |
| С                       | vector matrices of non linear functions of the continuous variables $x^N$ |
| Ci                      | molar concentration of component i  |
| Ci0                     | initial concentration of component i                                      |
| $CC_{k,0}$              | Capital Cost of the unit k with base size $Q_{k,0}$ , \$                  |
| ср                      | specific heat, $J/(kg K)$   |
| f                       | Annual fraction of the TIC  |
| $f_c$                   | cost factor   |
| $f_i$                   | Annual fraction of the TIC  |
| g                       | Working capital parameter   |
| $g_i$                   | parameter   |
| $K_k$                   | cost factor accounting for equipment -                                    |
| $m_k$                   | Relevant size factor of the unit k  |
| $n_h$                   | number of heat exchangers   |
| <i>n</i> <sub>int</sub> | number of subintervals  |
| $n_k$                   | number of outlet streams  |
| <b>N</b> comp           | number of components  |
| <i>n</i> <sub>d</sub>   | number of discrete values   |
| <i>n<sub>r</sub></i>    | number of reactions   |

 $\psi_{i,j}$ 

| <i>n</i> <sub>j</sub> | number of discrete values assumed for the non linear variable $x_j^N$   |
|-----------------------|---|
| $n_p$                 | number of products  |
| Ni                    | parameter in eq.19  |
| $P_{m,i}$             | permeate flux, $kg/(m^2 s)$   |
| $p_p$                 | unit market price, \$/kg  |
| r                     | Annual discount rate, %   |
| $R_i$                 | parameter in eq. 20   |
| $S_{k,0}$             | Base size of the unit k   |
| t                     | Tax rate, %   |
| Т                     | temperature, °C   |
| $T_j$                 | temperature of the subinterval $j$ °C                                   |
| $T_k^{IN}$            | temperatures of the inlet stream, $^{\circ}C$                           |
| $T_k^{OUT}$           | temperatures of the outlet stream, $^{\circ}C$                          |
| t <sub>ls</sub>       | Life span of the project, y   |
| $t_R$                 | Residence time for the vessels, h                                       |
| $U_k$                 | overall heat transfer coefficients in the heat exchanger, $W/(m^2 K)$   |
| $U_{nf}$              | linear flooding velocity in the distillation columns, m/s               |
| UB                    | upper bound for the flowrates   |
| Wi                    | mass concentration of component i                                       |
| Wi0                   | initial mass concentration of component i                               |
| $X_{B,i}$             | molar fraction of the bottom  |
| $x_{D,i}$             | molar fraction of the distillate  |
| α                     | vector of coefficients  |
| A                     | matrix of coefficients  |
| $\alpha_{lk/hk}$      | relative volatility between light key component and heavy key component |

| $lpha_{ij}$                  | discrete values of the function $a_j(x_j^N)$                               |
|------------------------------|--|
| β                            | vector of coefficients   |
| В                            | matrix of coefficients   |
| γ                            | vector of coefficients   |
| Г                            | matrix of coefficients   |
| δ                            | vector of coefficients   |
| Δ                            | matrix of coefficients   |
| $\Delta T$                   | mean temperature differences in the heat exchanger, $^{\circ}C$            |
| З                            | vector of coefficients   |
| З                            | fraction of the area available for vapour flow in the distillation columns |
| Ε                            | matrix of coefficients   |
| $\zeta_i$                    | vector of possible values of the variable                                  |
| $\zeta_{j,i}$                | discrete values assumed for the non linear variable $x_j^N$                |
| η <sub>ISO</sub>             | isoentropic efficiency   |
| <i><i><b>η</b>мес</i></i>    | mechanical efficiency  |
| $\overline{\eta}_{{}_{p,i}}$ | discrete values of the function $\eta_p(\tau)$                             |
| Н                            | matrix of coefficients   |
| $\lambda_i$                  | latent heat of the component i, J/kg                                       |
| $\mu_j^k$                    | discrete values of the split fraction for the stream k                     |
| $v_{i,j}$                    | stoichiometric coefficient for component i in reaction j                   |
| $\mathcal{V}_{j,l}$          | auxiliary parameter  |
| ρ                            | stream average density, kg/m <sup>3</sup>                                  |
| $ ho_{g}$                    | density of the vapour stream in the distillation columns, $kg/m^3$         |
| ξĭ                           | recovery of component i in a distillation column                           |
| τ                            | space time in the reactor, s   |

- $\bar{\tau}_i$  discrete values assumed for  $\tau$ , s
- $\omega_k$  auxiliary parameter

# Binary variables

| у          | binary variable  |
|------------|--|
| $y_j^k$    | binary variable for the selection of outlet stream flowrates |
| $y_i^d$    | binary variable introduced with the discretization           |
| $y_{ji}^d$ | binary variables introduced with the discretization          |

#### REFERENCES

- (1) Kamm, B.; Gruber, P. R.; Kamm, M. Biorefineries Industrial Processes and Products. In *Ullmann's Encyclopedia of Industrial Chemistry*; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2007.
- (2) FitzPatrick, M.; Champagne, P.; Cunningham, M. F.; Whitney, R. a. A Biorefinery Processing Perspective: Treatment of Lignocellulosic Materials for the Production of Value-Added Products. *Bioresour. Technol.* **2010**, *101* (23), 8915.
- (3) Balat, M.; Ayar, G. Biomass Energy in the World, Use of Biomass and Potential Trends. *Energy Sources* **2005**, *27* (10), 931.
- (4) Menon, V.; Rao, M. Trends in Bioconversion of Lignocellulose: Biofuels, Platform Chemicals & Biorefinery Concept. *Prog. Energy Combust. Sci.* **2012**, *38* (4), 522.
- (5) T. Werpy.; G. Petersen. Top Value Added Chemicals from Biomass Volume I Results of Screening for Potential Candidates from Sugars and Synthesis Gas Top Value Added Chemicals From Biomass Volume I: Results of Screening for Potential Candidates. **2004**, *1*.
- (6) E4tech; RE-CORD; WUR. From the Sugar Platform to Biofuels and Biochemicals. **2015**, contract No. ENER/C2/423.
- (7) Martín, M.; Grossmann, I. E. Optimal Simultaneous Production of I-Butene and Ethanol from Switchgrass. *Biomass and Bioenergy* **2014**, *61*, 93.
- (8) Poliakoff, M.; Licence, P. Sustainable Technology: Green Chemistry. *Nature* **2007**, *450* (7171), 810.
- (9) Geraili, A.; Salas, S.; Romagnoli, J. A. A Decision Support Tool for Optimal Design of Integrated Biorefineries under Strategic and Operational Level Uncertainties. *Ind. Eng. Chem. Res.* 2016, 55 (6), 1667.
- (10) Zondervan, E.; Nawaz, M.; de Haan, A. B.; Woodley, J. M.; Gani, R. Optimal Design of a Multi-Product Biorefinery System. *Comput. Chem. Eng.* **2011**, *35* (9), 1752.
- (11) Luo, L.; van der Voet, E.; Huppes, G. Biorefining of Lignocellulosic Feedstock Technical, Economic and Environmental Considerations. *Bioresour. Technol.* **2010**, *101* (13), 5023.
- (12) Kim, S. B.; Park, C.; Kim, S. W. Process Design and Evaluation of Production of Bioethanol and β-Lactam Antibiotic from Lignocellulosic Biomass. *Bioresour. Technol.* **2014**, *172*, 194.
- (13) Cheali, P.; Posada, J. a.; Gernaey, K. V.; Sin, G. Upgrading of Lignocellulosic Biorefinery to Value-Added Chemicals: Sustainability and Economics of Bioethanol-Derivatives. *Biomass and Bioenergy* **2015**, *75*, 282.
- (14) Cheali, P.; Quaglia, A.; Gernaey, K. V; Sin, G. Uncertainty Analysis in Raw Material and Utility Cost of Biorefinery Synthesis and Design. **2014**.
- (15) Floudas, C. a. Nonlinear and Mixed-Integer Optimization; Oxford, N. Y., PRESS, O. U., Eds.; 1995.
- (16) Duran, M. A.; Grossmann, I. E. Simultaneous Optimization and Heat Integration of Chemical Processes. *AIChE J.* **1986**, *32* (1), 123.
- (17) Westerlund, T.; Skrifvars, H.; Harjunkoski, I.; Pörn, R. An Extended Cutting Plane Method for a Class of Non-Convex MINLP Problems. *Comput. Chem. Eng.* **1998**, 22 (3), 357.
- (18) Gupta, Omprakash K., Ravindran, A. Branch and Bound Experiments in Convex Nonlinear

Integer Programming. 1985, 31 (12), 15333.

- (19) Trespalacios, F.; Grossmann, I. E. Review of Mixed-Integer Nonlinear and Generalized Disjunctive Programming Methods. *Chemie Ing. Tech.* **2014**, *86* (7), 991.
- (20) Ponce-Ortega, J. M.; Pham, V.; El-Halwagi, M. M.; El-Baz, A. a. A Disjunctive Programming Formulation for the Optimal Design of Biorefinery Configurations. *Ind. Eng. Chem. Res.* **2012**, *51*, 3381.
- (21) Gabriel, K. J.; El-Halwagi, M. M. Modeling and Optimization of a Bioethanol Production Facility. *Clean Technol. Environ. Policy* **2013**, *15* (6), 931.
- (22) Geraili, a.; Sharma, P.; Romagnoli, J. a. Technology Analysis of Integrated Biorefineries through Process Simulation and Hybrid Optimization. *Energy* **2014**, *73*, 145.
- (23) Bischi, A.; Taccari, L.; Martelli, E.; Amaldi, E.; Manzolini, G.; Silva, P.; Campanari, S.; Macchi, E. A Detailed MILP Optimization Model for Combined Cooling, Heat and Power System Operation Planning. *Energy* 2014, 74, 12.
- (24) Glover, F. Improved Linear Integer Programming Formulations of Nonlinear Integer Problems. *Manag. Sci.* **1975**, *22*, 455.
- (25) Kalitventzeff, B. Mixed Integer Non-Linear Programming and Its Application to the Management of Utility Networks. *Eng. Optim.* **1991**, *18* (1-3), 183.
- (26) Moro, L. F. L.; Pinto, J. M. Mixed-Integer Programming Approach for Short-Term Crude Oil Scheduling. *Ind. Eng. Chem. Res.* **2004**, *43* (1), 85.
- (27) Scott, F.; Venturini, F.; Aroca, G.; Conejeros, R. Selection of Process Alternatives for Lignocellulosic Bioethanol Production Using a MILP Approach. *Bioresour. Technol.* 2013, 148, 525.
- (28) Rackemann, D. W.; Doherty, W. O. The Conversion of Lignocellulosics to Levulinic Acid. *Biofuels, Bioprod. Biorefining* **2011**, *5* (2), 198.
- (29) Huang, H.-J.; Ramaswamy, S.; Al-Dajani, W. W.; Tschirner, U. Process Modeling and Analysis of Pulp Mill-Based Integrated Biorefinery with Hemicellulose Pre-Extraction for Ethanol Production: A Comparative Study. *Bioresour. Technol.* **2010**, *101* (2), 624.
- (30) Zimbardi, F.; Ricci, E.; Braccio, G. Technoeconomic Study on Steam Explosion Application in Biomass Processing. *Appl. Biochem. Biotechnol.* **2002**, *98-100*, 89.
- (31) Humbird, D.; Aden, A. Biochemical Production of Ethanol from Corn Stover : 2008 State of Technology Model Biochemical Production of Ethanol from Corn Stover. **2009**, No. August.
- (32) Mani, S.; Tabil, L. G.; Sokhansanj, S. Grinding Performance and Physical Properties of Wheat and Barley Straws, Corn Stover and Switchgrass. *Biomass and Bioenergy* 2004, 27, 339.
- (33) Eggeman, T.; Elander, R. T. Process and Economic Analysis of Pretreatment Technologies. *Bioresour. Technol.* 2005, *96* (18 SPEC. ISS.), 2019.
- (34) Mosier, N. Features of Promising Technologies for Pretreatment of Lignocellulosic Biomass. *Bioresour. Technol.* **2005**, *96* (6), 673.
- (35) Gabriel, K. J.; El-Halwagi, M. M. Modeling and Optimization of a Bioethanol Production Facility. *Clean Technol. Environ. Policy* **2013**, *15* (6), 931.
- (36) Chang, V. S.; Nagwani, M.; Holtzapple, M. T. Lime Pretreatment of Crop Residues Bagasse and Wheat Straw. *Appl. Biochem. Biotechnol.* **1998**, *74* (3), 135.

- (37) Pan, X.; Xie, D.; Yu, R. W.; Lam, D.; Saddler, J. N. Pretreatment of Lodgepole Pine Killed by Mountain Pine Beetle Using the Ethanol Organosolv Process: Fractionation and Process Optimization. *Ind. Eng. Chem. Res.* **2007**, *46* (8), 2609.
- (38) Lee, Y. Y.; Wu, Z.; Torget, R. W. Modeling of Countercurrent Shrinking-Bed Reactor in Dilute-Acid Total-Hydrolysis of Lignocellulosic Biomass. *Bioresour. Technol.* 2000, 71 (1), 29.
- (39) Girisuta, B.; Danon, B.; Manurung, R.; Janssen, L. P. B. M.; Heeres, H. J. Experimental and Kinetic Modelling Studies on the Acid-Catalysed Hydrolysis of the Water Hyacinth Plant to Levulinic Acid. *Bioresour. Technol.* **2008**, *99* (17), 8367.
- (40) Shen, J.; Wyman, C. E. Hydrochloric Acid-Catalyzed Levulinic Acid Formation from Cellulose: Data and Kinetic Model to Maximize Yields. *AIChE J.* **2012**, *58* (1), 236.
- (41) Ramaswamy, S.; Huang, H.-J.; Ramarao, B. V. Separation and Purification Technologies in *Biorefineries*; John Wiley & Sons, Ltd: Chichester, UK, 2013.
- (42) Dimian, A. C.; Bildea, C. S. *Chemical Process Design*; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2008.
- (43) Kadam, K. L.; Rydholm, E. C.; McMillan, J. D. Development and Validation of a Kinetic Model for Enzymatic Saccharification of Lignocellulosic Biomass. *Biotechnol. Prog.* 2004, 20 (3), 698.
- (44) Flores-Sánchez, A.; Flores-Tlacuahuac, A.; Pedraza-Segura, L. L. Model-Based Experimental Design to Estimate Kinetic Parameters of the Enzymatic Hydrolysis of Lignocellulose. *Ind. Eng. Chem. Res.* 2013, 52, 4834.
- (45) Leksawasdi, N.; Joachimsthal, E. L.; Rogers, P. L. Mathematical Modelling of Ethanol Production from Glucose/xylose Mixtures by Recombinant Zymomonas Mobilis. *Biotechnol. Lett.* **2001**, *23* (13), 1087.
- (46) Song, H.; Jang, S. H.; Park, J. M.; Lee, S. Y. Modeling of Batch Fermentation Kinetics for Succinic Acid Production by Mannheimia Succiniciproducens. *Biochem. Eng. J.* 2008, 40 (1), 107.
- (47) Farone, I. W. A.; John, E.; Ana, S. Method for the Production of Levulinic Acid and Its Derivatives. *US006054611A* **2000**, *36*.
- (48) O'Brien, D. J.; Roth, L. H.; McAloon, A. J. Ethanol Production by Continuous Fermentation–pervaporation: A Preliminary Economic Analysis. *J. Memb. Sci.* **2000**, *166* (1), 105.
- (49) Ayoub. Process for the Reactive Extractive Extraction of Levulinic Acid. US 7,378,549 B2, 2008.
- (50) Errico, M.; Rong, B. G.; Tola, G.; Spano, M. Optimal Synthesis of Distillation Systems for Bioethanol Separation. Part 1. Extractive Distillation with Simple Columns. *Ind. Eng. Chem. Res.* 2013, 52 (4), 1612.
- (51) Kwiatkowski, J. R.; McAloon, A. J.; Taylor, F.; Johnston, D. B. Modeling the Process and Costs of Fuel Ethanol Production by the Corn Dry-Grind Process. *Ind. Crops Prod.* 2006, 23 (3), 288.
- (52) Alvarez, M. E. T.; Moraes, E. B.; Araujo, W. A.; Maciel Filho, R.; Wolf-Maciel, M. R. Development of a Mathematical Model for Studying Bioethanol–water Separation Using Hydrophilic Polyetherimide Membrane. *J. Appl. Polym. Sci.* **2008**, *107* (4), 2256.

- (53) Yedur, S.; Berglund, K. A.; Dilum D.; Dunuwila. Succinic Acid Production and Purification. US 6,265,190 B1, 2001.
- (54) Song, H.; Huh, Y. S.; Lee, S. Y.; Hong, W. H.; Hong, Y. K. Recovery of Succinic Acid Produced by Fermentation of a Metabolically Engineered Mannheimia Succiniciproducens Strain. J. Biotechnol. 2007, 132 (4), 445.
- (55) Huh, Y. S.; Jun, Y. S.; Hong, Y. K.; Song, H.; Lee, S. Y.; Hong, W. H. Effective Purification of Succinic Acid from Fermentation Broth Produced by Mannheimia Succiniciproducens. *Process Biochem.* **2006**, *41*, 1461.
- (56) Hamelinck, C. N.; Van Hooijdonk, G.; Faaij, a. P. C. Ethanol from Lignocellulosic Biomass: Techno-Economic Performance in Short-, Middle- and Long-Term. *Biomass and Bioenergy* 2005, 28 (4), 384.
- (57) Aden, a; Ruth, M.; Ibsen, K.; Jechura, J.; Neeves, K.; Sheehan, J.; Wallace, B.; Montague, L.; Slayton, a; Lukas, J. Lignocellulosic Biomass to Ethanol Process Design and Economics Utilizing Co-Current Dilute Acid Prehydrolysis and Enzymatic Hydrolysis for Corn Stover. *Other Inf. PBD 1 Jun 2002* 2002, No. June, Medium: ED; Size: 154 pages.
- (58) Cardona, C. A.; Sanchez, O. J. Fuel Ethanol Production: Process Design Trends and Integration Opportunities. *Bioresour. Technol.* **2007**, *98* (12), 2415.
- (59) Sun, Y.; Cheng, J. Hydrolysis of Lignocellulosic Materials for Ethanol Production: A Review Q. *Bioresour. Technol.* 2002, 83 (1), 1.
- (60) Biegler L.T., Grossmann I.E., W. A. W. Systematic Methods of Chemical Process Design; Prentice-Hall, 1997.
- (61) Linnhoff, B.; Hindmarsh, E. The Pinch Design Method for Heat Exchanger Networks. *Chem. Eng. Sci.* **1983**, *38* (5), 745.
- (62) Vogelpohl, A. Die Naeherungsweise Berechnung Der Rektifikation von Gemischen Mit Binaeren Azeotropen. *Chem. Ing. Tech.* **1974**, *46* (5), 195.
- (63) Wang, B.; Gebreslassie, B. H.; You, F. Sustainable Design and Synthesis of Hydrocarbon Biorefinery via Gasification Pathway: Integrated Life Cycle Assessment and Technoeconomic Analysis with Multiobjective Superstructure Optimization. *Comput. Chem. Eng.* 2013, 52, 55.
- (64) Kazi, F. K.; Fortman, J.; Anex, R. Techno-Economic Analysis of Biochemical Scenarios for Production of Cellulosic Ethanol. *Natl. Renew. Energy Lab.* **2010**, No. June, 102.
- (65) Pintarič, Z.; Kravanja, Z. Selection of the Economic Objective Function for the Optimization of Process Flow Sheets. *Ind. Eng. Chem. Res.* **2006**, 4222.
- (66) Kasaš, M.; Kravanja, Z.; Novak Pintarič, Z. Suitable Modeling for Process Flow Sheet Optimization Using the Correct Economic Criterion. *Ind. Eng. Chem. Res.* **2011**, *50*, 3356.
- (67) Mosier, N. S.; Hendrickson, R.; Brewer, M.; Ho, N.; Sedlak, M.; Dreshel, R.; Welch, G.; Dien, B. S.; Aden, A.; Ladisch, M. R. Industrial Scale-up of pH-Controlled Liquid Hot Water Pretreatment of Corn Fiber for Fuel Ethanol Production. *Appl. Biochem. Biotechnol.* 2005, 125 (2), 77.
- (68) Teymouri, F.; Laureano-Perez, L.; Alizadeh, H.; Dale, B. E. Optimization of the Ammonia Fiber Explosion (AFEX) Treatment Parameters for Enzymatic Hydrolysis of Corn Stover. *Bioresour. Technol.* 2005, 96 (18 SPEC. ISS.), 2014.
- (69) Girisuta, B.; Janssen, L. P. B. M.; Heeres, H. J. Kinetic Study on the Acid-Catalyzed

Hydrolysis of Cellulose to Levulinic Acid. Ind. Eng. Chem. Res. 2007, 46 (6), 1696.

(70) Song, H.; Jang, S. H.; Park, J. M.; Lee, S. Y. Modeling of Batch Fermentation Kinetics for Succinic Acid Production by Mannheimia Succiniciproducens. *Biochem. Eng. J.* 2008, 40 (1), 107.

#### **Table captions**

Table 1: Pretreatment parameters for the optimization model.

Table 3: Parameter and residence time ranges for the reactors.

Table 2: Product yields for the purification sections.

Table 4: Biomass composition and main parameters for the optimization model.

Table 5: Sensitivity results of the economic analysis on plant size and biochemicals selling price by maximizing the net present value.

# **Figure captions**

- *Figure 1: Superstructure of the multi-product lignocellulosic biorefinery.*
- Figure 2: Reaction network of dilute acid hydrolysis of cellulose and hemicellulose.
- Figure 3: Reaction network of dilute acid hydrolysis of hemicellulose.
- *Figure 4: Reaction network of co-fermentation of ethanol.*
- *Figure 5: Reaction network of fermentation to produce succinic acid.*
- Figure 6: Reactions to convert succinic acid to diammonium succinate (a) and to obtain succinic acid as a solid precipitate product (b).
- Figure 7: Optimal NPV flowsheets of biorefinery for base case and variants: a) base case; b) biomass feed rate 5 t/h (0.1 times the base case); c) levulinc acid selling price7.50 \$/kg(+50% of the base case).
- *Figure 8: Capital cost for the lignocellulosic biorefinery by maximizing the net present value and the internal rate of return.Black: pretreatment section; dark grey: hydrolysis and fermentation section; light grey: separation and purification section; striped grey: thermochemical conversion section.*

- *Figure 9:* Cost distribution for the lignocellulosic biorefinery by maximizing the net present value and the internal rate of return.Black:raw material cost; dark grey: enzymes and reactants cost; light grey: utility cost; striped grey: other manufacturing cost.
- *Figure 10: Objective function (net present value) and internal rate of return for each value of the discount rate. Net present value (•); internal rate of return (■).*
- Figure 11: Biomass allocation (a) and product yields (b) for the base case by maximizing the net present value and the internal rate of return.Black: levulinic acid; dark grey: ethanol; light grey: succinic acid.

# TABLE OF CONTENTS GRAPHIC



| Pretreatment                   | Temperature<br>(°C) | Residence time<br>(min) | Concentration                        |
|--------------------------------|---------------------|-------------------------|--------------------------------------|
| Steam explosion <sup>30</sup>  | 220 °C              | 3                       | /                                    |
| Liquid hot water <sup>67</sup> | 190 °C              | 15                      | /                                    |
| AFEX <sup>68</sup>             | 90 °C               | 5                       | 50% wt                               |
| Dilute acid <sup>31</sup>      | 160 °C              | 10                      | 1% wt H <sub>2</sub> SO <sub>4</sub> |
| Lime <sup>36</sup>             | 120 °C              | 120                     | 9% Ca(OH)2 w/w dry biomass           |
| Organo solvent <sup>37</sup>   | 180 °C              | 60                      | 60% ethanol w/w dry biomass          |

Table 1: Pretreatment parameters for the optimization model.

Table 2: Product yields for the purification sections.

| Equipment  | Main product yield (%wt) |
|--|--------------------------|
| Chromatographic column to purify levulinic acid <sup>47</sup>            | 95                       |
| Distillation column to purify levulinic acid <sup>47,40</sup>            | 95                       |
| Reactive solvent extraction column and distillation column to purify     | 84                       |
| levulinic acid <sup>49</sup>   |                          |
| Distillation columns to ethanol <sup>42</sup>                            | 99                       |
| Pervaporation membranes to ethanol <sup>48,52</sup>                      | 95                       |
| Distillation column with entrainer <sup>50</sup>                         | 95                       |
| Molecular sieves to ethanol <sup>51</sup>                                | 90                       |
| Reactive crystallization of succinic acid <sup>53</sup>                  | 99                       |
| Extraction and distillation column to purify succinic acid <sup>54</sup> | 90                       |

| Table 3: Parameter and residence time ranges for the reactors. |
|--|
|--|

| Reactor  | Temperature | Residence | Yield                  | Reactor |
|--|-------------|-----------|------------------------|---------|
|  | (°C)        | time      | (kgproduct/kgreactant) | type    |
| Reactor to convert                                 | 150         | 0-120 min | Glucose: 0-30%         | PFR     |
| cellulose/glucose to                               |             |           | Xylose: 0-60%          |         |
| levulinic acid with H <sub>2</sub> SO <sub>4</sub> |             |           | Levulinic acid: 0-55%  |         |
| as catalyst <sup>69</sup>                          |             |           |                        |         |
| Reactor to convert                                 | 150         | 0-50 min  | Glucose: 0-15%         | PFR     |
| cellulose/glucose to                               |             |           | Xylose: 0-30%          |         |
| levulinic acid HCl as                              |             |           | Levulinic acid: 0-35%  |         |
| catalyst <sup>40</sup>                             |             |           |                        |         |
| Enzymatic reactor to                               | 45          | 0-72 h    | Glucose: 0-70%         | CSTR    |
| produce glucose and                                |             |           | Xylose: 0-60%          |         |
| xylose <sup>43,44</sup>                            |             |           |                        |         |
| Reactor to convert                                 | 150         | 0-120 min | Xylose : 0-70%         | PFR     |
| hemicellulose to xylose <sup>38</sup>              |             |           |                        |         |
| Fermenter to produce                               | 30          | 0-30 h    | Ethanol: 0-47%         | Batch   |
| ethanol <sup>45</sup>                              |             |           |                        |         |
| Fermenter to produce                               | 40          | 0-15 h    | Succinic acid: 0-47%   | Batch   |
| succinic acid <sup>70</sup>                        |             |           |                        |         |

| Biomass                      | Hardwood |  |  |  |
|------------------------------|----------|--|--|--|
| Plant Life (y)               | 20       |  |  |  |
| Feedstock (t/h)              | 50       |  |  |  |
| Cellulose (%dry)             | 50       |  |  |  |
| Hemicellulose (%dry)         | 16       |  |  |  |
| Lignin (%dry)                | 28       |  |  |  |
| Other Compounds (%dry)       | 6        |  |  |  |
| Levulinic Acid Price (\$/kg) | 5.00     |  |  |  |
| Succinic Acid Price (\$/kg)  | 7.50     |  |  |  |
| Ethanol Price (\$/kg)        | 0.75     |  |  |  |
| Biomass Price (\$/t)         | 40       |  |  |  |
| Discount Rate (%)            | 8.00     |  |  |  |
| Enzyme cost (\$/kg)          | 5.00     |  |  |  |
| $H_2SO_4 \cos(kg)$           | 0.08     |  |  |  |
| HCl cost (\$/kg)             | 0.30     |  |  |  |
| Ammonia cost (\$/kg)         | 0.40     |  |  |  |
| Ca(OH) <sub>2</sub> (\$/kg)  | 0.06     |  |  |  |
| Electricity cost (\$/MWhe)   | 150      |  |  |  |
| Octanol cost (\$/kg)         | 5.00     |  |  |  |
| Taxes (%)                    | 40       |  |  |  |
| $\Delta T$ (°C)              | 40       |  |  |  |
| $\Delta T_{MIN}$ (°C)        | 20       |  |  |  |

Table 4: Main parameters for the optimization model.

|                                | Base<br>case | Effect of plant size |      | Effect of<br>succinic acid<br>price |       | Effect of levulinic acid price |      |
|--------------------------------|--------------|----------------------|------|-------------------------------------|-------|--------------------------------|------|
| Biomass feedstock<br>rate(t/h) | 50           | 15                   | 150  | 50                                  | 50    | 50                             | 50   |
| Succinic acid Price (\$/kg)    | 7.50         | 7.50                 | 7.50 | 3.75                                | 11.25 | 7.50                           | 7.50 |
| Levulinic acid Price (\$/kg)   | 5.00         | 5.00                 | 5.00 | 5.00                                | 5.00  | 2.50                           | 7.50 |
| Biomass allocation to          |              |                      |      |                                     |       |                                |      |
| Levulinic acid (%)             | 46           | 0                    | 46   | 90                                  | 0     | 0                              | 86   |
| Ethanol (%)                    | 12           | 26                   | 12   | 4                                   | 26    | 26                             | 5    |
| Succinic acid (%)              | 42           | 74                   | 42   | 6                                   | 74    | 74                             | 9    |
| Yield to                       |              |                      |      |                                     |       |                                |      |
| Levulinic acid (%)             | 8.9          | 0.0                  | 9.3  | 13.4                                | 0     | 0                              | 22.2 |
| Ethanol (%)                    | 2.2          | 3.9                  | 2.1  | 1.0                                 | 3.9   | 3.9                            | 1.3  |
| Succinic acid (%)              | 5.7          | 10.9                 | 6.2  | 1.2                                 | 10.0  | 10.0                           | 2.3  |
| Total (%)                      | 16.7         | 14.7                 | 17.6 | 15.6                                | 13.9  | 13.9                           | 25.9 |
| Total Investment Cost (M\$)    | 413          | 103                  | 1088 | 338                                 | 249   | 231                            | 1319 |
| Total Annual Cost (M\$/y)      | 118          | 44                   | 328  | 125                                 | 120   | 122                            | 236  |
| Rev (M\$/y)                    | 319          | 91                   | 1022 | 259                                 | 416   | 280                            | 666  |
| IRR (%)                        | 25           | 22                   | 29   | 20                                  | 47    | 31                             | 16   |
| NPV (M\$)                      | 620          | 142                  | 2423 | 358                                 | 1215  | 570                            | 957  |

Table 5: Sensitivity results of the economic analysis on plant size and biochemicals selling price by maximizing the net present value



Figure 1: Superstructure of the multi-product lignocellulosic biorefinery.

 $\begin{array}{ccc} \text{Cellulose} \rightarrow & \text{Glucose} \rightarrow \text{HMF} \rightarrow \text{Levulinic Acid} + \text{Formic Acid} \\ & & \downarrow \\ \text{Decomposed} & \text{Decomposed} \\ \text{Compounds} & \text{Compounds} \\ \text{Hemicellulose} \rightarrow & \text{Xylose} \rightarrow \text{Furfural} \\ & & \downarrow \\ \text{Decomposed} & \text{Decomposed} \\ \text{Compounds} & \text{Compounds} \end{array}$ 

Figure 2: Reaction network of dilute acid hydrolysis of cellulose and hemicellulose.

 $\begin{array}{ccc} \text{Hemicellulose} \to \text{Xylose} & \to \text{Furfural} \\ & & \downarrow \\ \text{Decomposed} & \text{Decomposed} \\ \text{Compounds} & \text{Compounds} \end{array}$ 

Figure 3: Reaction network of dilute acid hydrolysis of hemicellulose.

Glucose  $\rightarrow$ 2 Ethanol+2 CO<sub>2</sub>

3 Xylose  $\rightarrow$ 5 Ethanol+5 CO<sub>2</sub>

Figure 4: Reaction network of co-fermentation of ethanol.

Glucose+2 CO<sub>2</sub>  $\rightarrow$ 2 Succinic Acid+ O<sub>2</sub>

3 Xylose+5 CO<sub>2</sub>  $\rightarrow$ 5 Succinic Acid+2.5 O<sub>2</sub>

Glucose  $\rightarrow$  3 Acetic Acid

Glucose  $\rightarrow$  2 Lactic Acid

2 Xylose  $\rightarrow$  5 Acetic acid

3 Xylose  $\rightarrow$  5 Lactic acid

Figure 5: Reaction network of fermentation to produce succinic acid.

Succinic Acid +2 NH<sub>3</sub>  $\rightarrow$  Diammonium Succinate

(a)

Diammonium Succinate +2 NH<sub>4</sub> HSO<sub>4</sub>  $\rightarrow$  Succinic Acid + 2 (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> (b)

Figure 6: Reactions to convert succinic acid to diammonium succinate (a) and to obtain succinic acid as a solid precipitate product (b).



Figure 7: Optimal NPV flowsheets of biorefinery for base case and variants: a) base case; b) biomass feed rate 5 t/h (0.1 times the base case); c) levulinc acid selling price7.50  $\frac{1}{kg}(+50\%)$  of the base case).



Figure 8: Capital cost for the lignocellulosic biorefinery by maximizing the net present value and the internal rate of return.Black: pretreatment section; dark grey: hydrolysis and fermentation section; light grey: separation and purification section; striped grey: thermochemical conversion section.



*Figure 9:* Cost distribution for the lignocellulosic biorefinery by maximizing the net present value and the internal rate of return.Black:raw material cost; dark grey: enzymes and reactants cost; light grey: utility cost; striped grey: other manufacturing cost.



*Figure 10: Objective function (net present value) and internal rate of return for each value of the discount rate. Net present value (●); internal rate of return (■).* 



Figure 11: Biomass allocation (a) and product yields (b) for the base case by maximizing the net present value and the internal rate of return.Black: levulinic acid; dark grey: ethanol; light grey: succinic acid.