

# Modelling of Biomass-to-X: Challenges and Strategies

Matteo Gilardi<sup>a</sup>, Filippo Bisotti<sup>a</sup>, Olaf T. Berglihn<sup>a</sup>, Theresa Rücker<sup>a</sup>, Thomas Brück<sup>b</sup>, Bernd Wittgens<sup>a,\*</sup>

<sup>a</sup> SINTEF Industry – P.O. Box 4760 Torgarden, NO-7465 Trondheim, Norway

<sup>b</sup> Technical University of Munich – Lichtenbergstr. 4, 85748 Garching, Germany

[bernd.wittgens@sintef.no](mailto:bernd.wittgens@sintef.no)

The progressive request for replacing fossil fuels and net-zero emissions by 2050 is forcing a change in the current paradigm and economy. Bioprocesses offer a significant alternative to conventional refineries, as they convert biomass into valuable platform chemicals, fuels, and/or heat and energy. The concept is disruptive, but the modelling is not straightforward due to the intrinsic complexity of the feedstock, the unconventional transformation (i.e., pre-treatment and chemical/thermal conversion, and product purification), and the number of species involved in the system. These aspects are of paramount importance and represent significant hurdles for modelling. In the present work, we review the main strategies used for bioprocesses modelling by highlighting advantages and drawbacks. Finally, the work presents a few applications applied in process design and preliminary assessment for biomass-to-X projects.

## 1. Introduction

The urgent shift from a fossil- to a bio-based economy for the replacement of fossil fuels and fossil carbon-based chemicals opens to new solutions. Biomass-to-X offers a sustainable route to produce high-added-value products, heat, and energy. The development of models for bio-based processes is essential for design purposes, energy assessments, and cost estimates. However, modelling and design are not straightforward owing to the intrinsic complexity and variability of the feedstocks and their transformation. Thus, it requires simplified, but still accurate approaches to ensure good predictions, while keeping reasonable computational effort. This work is divided into two main sections. The first part reviews the state-of-the-art of strategies and approaches for the modelling of biomass conversion processes by highlighting the benefits and drawbacks of short-cut, data-driven, and rigorous models (Martin et al., 2023). This section introduces the importance of simulations for designing, upscaling from laboratory scale (batch or semi-batch apparatus) to pilot/industrial facilities (continuous operation), and the potential for machine learning to generate sub-models for bioprocesses (Valedandi et al., 2023). The second part deals with the application of the mentioned modelling strategies for the sake of process design, upscaling, unit operation sizing, and techno-economic assessment (TEA) in two specific projects: VALUABLE project, a Horizon Europe Research and Innovation project (VALUABLE webpage) and Bio4Fuels, an FME Norwegian project (Bio4Fuels webpage). For both projects, dedicated sub-models for each main unit operation have been developed as plug-ins and then inter-connected to build a comprehensive process flow using COCO-COFE v3.6, a license-free CAPE-open simulation software developed by AmsterChem. The developed framework is then exploited to maximize the plant productivity through specific sensitivity analysis, accounting also for energy integration and by-product stream recycling when possible. Mass and energy balances from the simulations are used as a basis framework for unit sizing, investment, raw materials, and utility costs estimation, as well as for the quantification of pollutant emissions and life cycle assessment, which are not part of the present work.

## 2. Modelling strategies and approaches

Process simulation is a key step for the design, scale-up, and implementation of bioprocesses and biorefineries to address the energy demand and material balances of the entire process. For a preliminary upscaling, the

process configuration and residence times tested on a lab scale can be maintained, while consumables and unit sizes are re-defined based on the mass flow ratio between the final and the original scale. Moreover, the simulation provides background information for the TEA and life-cycle assessment (LCA), and it optimizes or gets energy-integrated plants to lower the demand for external utilities such as cooling water and natural gas. As mentioned, the complexity and large variety of the feedstocks are reflected in a wide plethora of products generated during biomass pre-treatment and conversion. Moreover, many unconventional units to treat the biomass are a further complication as well as the presence of complex reaction schemes, mass transfer limitations, and thermodynamics. For the first hurdle, it is possible to account for a lumping approach, where a limited number of species to track the evolution of the biomass treatment and cover all possible chemical classes (e.g., phenols, ether, ketones, etc.), is considered without loss of generality and accuracy (Bisotti et al, 2023; Gilardi et al, 2023). Since flowsheets can only handle a limited complexity in reactor design, systems can be simplified by accounting only for those reactions having a higher impact on the product distribution and overall reaction rate (e.g., definition of the Rate Determining Step). Moreover, we adopt smart approaches like the Enhancement Factor method to describe mass transfer effects within a given column or reactor. In our perspective, depicted in Figure 1, each module/unit operation is defined in a sub-model. These are added as plug-ins and integrated within the process simulator to provide a single flowsheet, whose results are finally exploited for TEA and LCA. Figure 1 summarizes the tools adopted to perform optimization, LCA, and TEA. The following sub-sections deal with the modelling and strategies which are currently applied for bioprocess and biorefinery design, and we provide an overview of the most promising options.

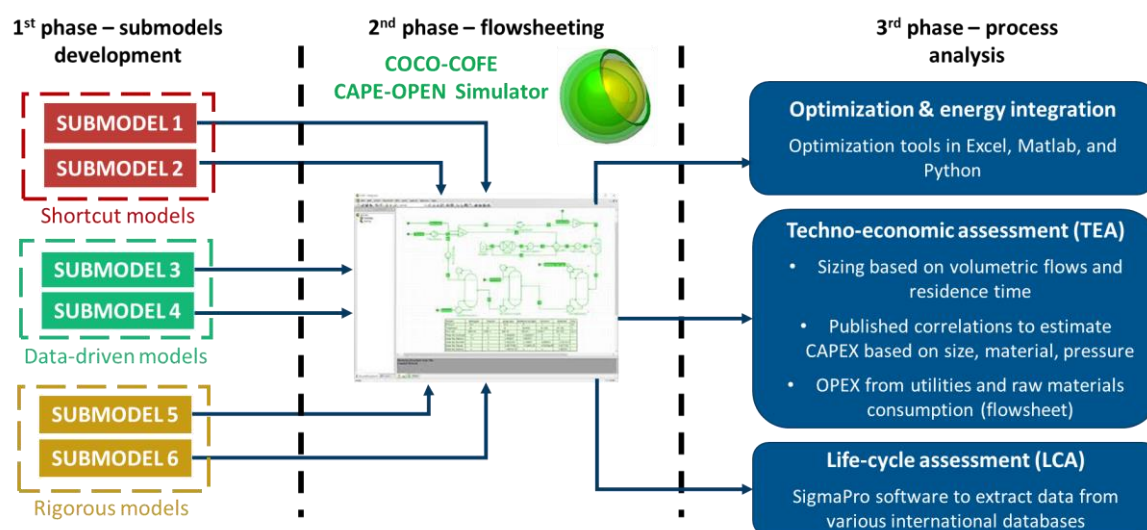


Figure 1: Overview of strategy for bioprocesses modelling and assessment.

## 2.1 Shortcut models

Shortcut models are at the highest level of simplification. They are designed to get preliminary mass and energy balance, using basic thermodynamics and assumptions. This simplified approach accounts for fixed parameters (generally conversion or product yield) derived from experience, existing facilities, and literature. Complex units are modelled as a black box, whenever there is neither the need to provide details on the process nor enough data available to generate comprehensive/detailed models. Short-cut models aim to address preliminary estimates of the process without any need for testing the sensitivity to the operating condition and inputs. Shortcut models are not suitable for scale-up as they generally do not apply to a large domain. Shortcut models are simplified input-output models that do not include a description of the internal functionality.

## 2.2 Empirical and data-driven models

Empirical models rely on observations rather than on physically rigorous relationships developed for the assigned system. These are often yield or conversion-based models, where the yield (or any relevant key performance indicator) of a process is defined as a function of key operating conditions using adjustable parameters. These are often tuned upon measurements from dedicated experimental campaigns covering the operating conditions domain of interest. Needless to mention, such models are simple, flexible, and easy to use.

Their main disadvantage is that they cannot be extrapolated outside the tested range of operating conditions and feedstocks. Data-driven models are a natural evolution of empirical models, where, nowadays, machine learning techniques find more applications such as the design of experiments (Martin and Grossmann, 2012; Quirino et al, 2022), artificial neural networks (Henao and Maravelias, 2011, Galeazzi et al, 2024), or surrogate models (Caballero and Grossmann, 2008; Di Pretoro et al, 2021).

### 2.3 Rigorous (first-principle) models

First-principle models follow a rigorous theoretical approach. This approach is general and considers all possible events occurring within the system through detailed kinetic schemes, whenever needed, mass and heat transfer, momentum balance as well as any chemical/physical phenomena relevant to characterize the system. The development of rigorous models is time-consuming and requires large datasets for validation. Despite the mentioned drawbacks, their validity is general, and they are applied in many fields because they are not tailored to a specific application (Dente et al, 2007).

## 3. Examples

We are providing some examples of how the mentioned approaches have been applied to different projects.

### 3.1 Valuable

VALUABLE is project funded by the European Union Horizon Europe Research and Innovation Action program targeting the valorization of *Aspergillus niger* (AN) fungal biomass, a byproduct from citric acid production, for the sustainable production of triglyceride-rich yeast oils. Additionally, the focus is on the recovery of chitin-rich waste to produce non-animal derived chitosan with applications in cosmetics and bio-based chemical markets (VALUABLE webpage, 2022). A block flow diagram of the process is shown in Figure 2. The enzymatic solubilization of AN results in a sugar-rich liquor and a chitin-rich solid fraction. Chitin is conveyed to a bioreactor where it is processed into chitosan, while sugar is added as a nutrient in a fermentation chamber to grow Oleaginous Yeasts cell mass and strain. The grown biomass is described as triglycerides and non-lipid biomass (NLB). NLB is hydrolyzed producing sugars and amino acids, also releasing some dead cell-mass. The products are finally separated using a gravimetric force.

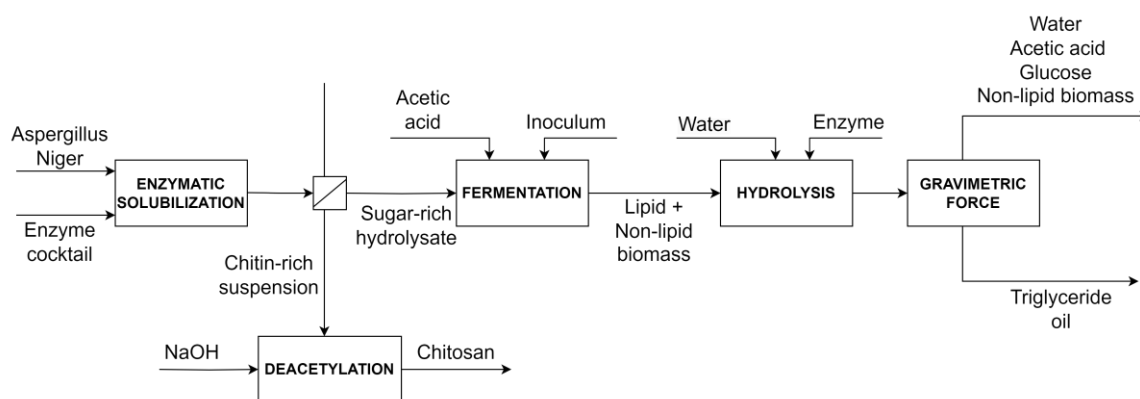
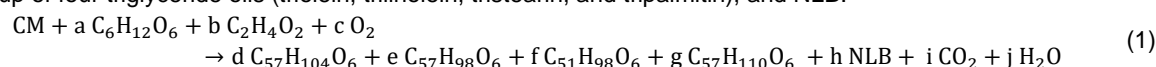


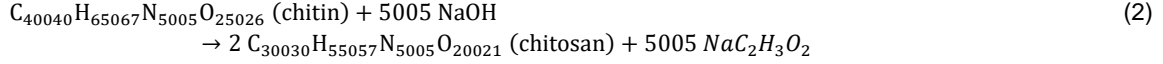
Figure 2: Block flow diagram of *Aspergillus niger* valorization process modelled in VALUABLE project.

SINTEF took advantage of laboratory-scale experimental data for the single conversion steps (e.g., enzymatic solubilization, fermentation to grow yeast cell mass, hydrolysis to release the triglyceride-rich oil, chitin conversion into chitosan) to develop a process flowsheet for upscaling the proposed process to an industrial scale. Given the complexity of the chemistry involved in this bioprocess (e.g., the complex molecular structure of enzymes and several elementary steps involved in the enzymatic processes), data-driven conversion models are adopted to deal with the sub-units. A first application of data-driven models in VALUABLE is the characterization of aerobic fermentation (reaction (1)). The inoculum grows by consuming glucose and acetic acid as an additional carbon source. The resulting grown yeast cell mass (CM) is defined as lipid biomass made up of four triglyceride oils (triolein, trilinolein, tristearin, and tripalmitin), and NLB.



The optimal set of stoichiometric coefficients (a to j) is regressed by multivariable optimization to minimize the deviations in the atomic balance for the single elements under the following constraints (C1-3) deriving from

experimental observations: (C1) the co-carbon sourcing from sugar and acetic acid is assigned, (C2) the obtained lipid cell mass is four times the yielded NLB, and (C3) the relative composition of lipid biomass is 45% triolein, 28% tristearin, 22% tripalmitin, and 5% trilinolein. The obtained sub-model is implemented in COCO-COFE and integrated with the other sub-models to simulate the entire process. The conversion of reaction (4) is tuned to achieve the desired triglyceride oil productivity for the up-scaled plant, equal to 0.6 ton/ton of AN. Another example is the sub-model for chitin conversion into chitosan through a chemical route (NaOH treatment). A representative chemical reaction is considered (2), in compliance with atomic mass-balance requirements.



For simplicity, a chemical formula for chitin is defined based on in-house measurements of the degree of polymerization of the input to the deacetylation reactor. The degree of freedom of the system, namely the conversion, is tailored to match the following requirement: the degree of deacetylation (i.e., the fraction of deacetylated chitosan in the monomeric chitin-chitosan mixture resulting from deacetylation) predicted by the model must be compliant with minimum deacetylation degree required for chitosan production, which is equal to 50%.

### 3.2 Bio4Fuels

Bio4Fuels is part of the FME projects funded by the Norwegian Research Council (Bio4Fuel webpage, 2018). The project investigates the conversion of spruce wood chips into fermentable sugars (Gilardi et al., 2023) and the valorization of unconverted lignin residue into additional bio-oil through pyrolysis (Bisotti et al., 2023). SINTEF developed a plausible industrial process (i.e., preliminary process flow diagram, sub-models, and full flowsheet, as shown in Figure 3) based on experimental laboratory scale data generated within the project (Hansen et al, 2022). The flowsheet interlinks several sub-models (e.g., enzymatic saccharification, steam explosion, fermentation, etc.) developed according to different strategies, as described in Section 2. Examples of shortcut (fermentation), data-driven (enzymatic saccharification) and rigorous (pyrolysis) sub-models follow.

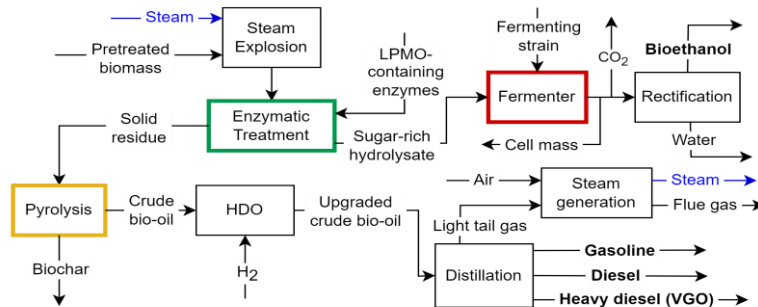


Figure 3: Bio4Fuels block flow diagram and implemented sub-models (highlighted processes are described in the work).

The fermentation to convert sugar-rich hydrolysate into bioethanol is designed as a black box according to the guidelines of the National Renewable Energy Laboratory, NREL (Humbird, 2011). The retrieved pieces of information include the operating temperature and residence time, glucose and xylose conversions, as well as the yeast consumption and the amount of fermentation strain. The simulation re-scales the fermenter to the size of interest for the actual bioethanol productivity in Bio4Fuels. In other words, we used details (i.e., key performance indicators) from an existing facility to re-design the same process in another context.

For enzymatic saccharification, a data-driven soft model is tuned to experimental data on glucose yield ( $\gamma$ ) under different operating conditions published by Hansen (Hansen et al., 2022). The model accounts for sugar yield as a function of residence time ( $t$ ), enzyme concentration ( $[E]$ ), and steam explosion temperature ( $T_{SE}$ ), as highlighted in expressions (5) to (8).

$$\gamma(t) = \gamma(0) + h \cdot [1 - \exp(-k_R \cdot t)] \quad (5)$$

$$\gamma(0) = a_1 \cdot [E] + a_2 \quad (6)$$

$$h(T_{SE}, [E]) = (a_3 \cdot T_{SE} + a_4) \cdot [E] + (a_5 \cdot T_{SE} + a_6) \quad (7)$$

$$k_R(T_{SE}, [E]) = (a_7 \cdot T_{SE} + a_8) \cdot [E] + (a_9 \cdot T_{SE} + a_{10}) \quad (8)$$

Coefficients  $a_1$  to  $a_{10}$  are obtained because of a multi-variable optimization, where the target is the minimization of the discrepancy between the glucose yield predicted by the model and the corresponding experimental observations. The charts (Figure 4) show a graphical comparison between the data-driven model and the data exploited for its fitting in terms of saccharification yield as a function of the residence time at different enzyme concentrations and steam explosion temperature.

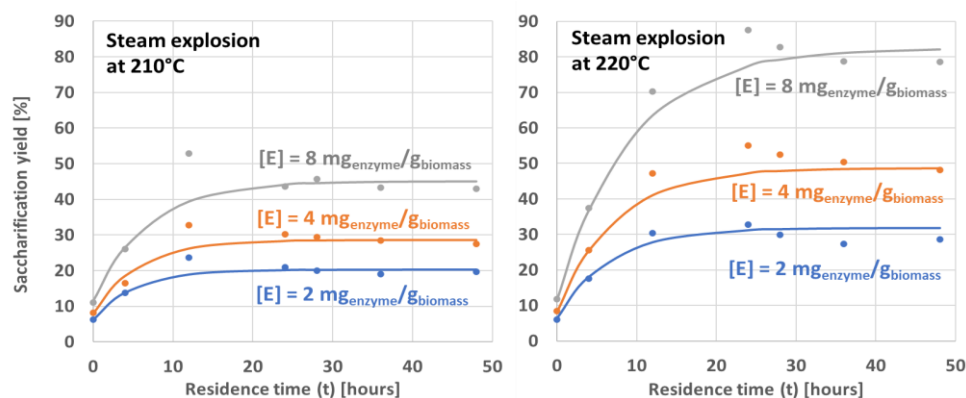


Figure 4: Enzymatic saccharification yield as a function of residence time at different enzyme concentrations  $[E]$  and steam explosion temperature ( $T_{SE}$ ): experimental data ( $\bullet$ ) and model prediction (-).

Conversely, the pyrolysis chamber has been designed based on rigorous models retrieved from the literature. Indeed, the design of the reactor and the kinetics of conversion reactions are included to characterise the evolution of the system composition as a function of the residence time (along the reactor coordinate). The chamber is modelled in Python as a fluidized bed reactor operated isothermally at 550 °C. Given the complexity of the system, a lumped kinetic is adopted. The kinetics for cellulose and hemicellulose decomposition are retrieved from Ranzi (2008), while the updated model by Dussan (2019) is implemented for lignin pyrolysis. A mixed model is preferred considering that Dussan's model benefits compared to Ranzi's model for lignin degradation since it can better predict the light gas and char fractions and more chemical classes are included. Both Dussan and Ranzi models have been validated over several thermogravimetric measurements for different kinds of biomass. The solid and gas phases rise along the column while the solid is gradually degraded to lighter chemicals. The details on the model and the pyrolysis reactor are available in Bisotti et al. (2023). The model is exploited to define the optimal operating temperature to achieve a reasonable lignin degree of conversion and to quantify char and light gas formation, as well as how the product distribution in the bio-oil changes at different temperatures. Results are shown in Figure 5.

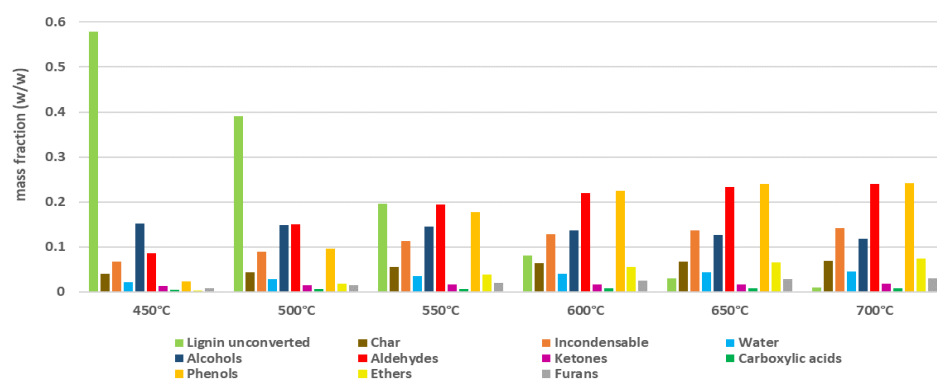


Figure 5: Sensitivity analysis on pyrolysis operating temperature and product distribution as predicted by the lumped kinetic model exploited in Bio4Fuels.

The minimum required temperature to convert at least 90% of lignin is 600°C. Moreover, phenols and ether formation considerably increase at higher temperatures, while the alcohol fraction drops in line with thermodynamic stability considerations. The aforementioned sub-models have been integrated into a single flowsheet, whose results will be discussed in other works.

#### 4. Conclusions

A review of the main strategies and approaches for modelling, scale-up, optimization, and techno-economic assessment of bioprocesses is proposed, and we highlighted the main challenges related to the complex nature of feedstocks and products as well as unconventional units involved in the biomass treatment. This contribution showed how simplified (i.e., shortcut and data-driven) and rigorous models are exploited for the characterisation of single-unit operation and then effectively combined into an integrated process simulation. Modelling strategies must be tuned to reach the characterization level of detail required for a specific process while maintaining a reasonable computational effort. This work describes how the mentioned approaches are applied and integrated into the specific context of two projects (i.e., VALUABLE and Bio4Fuels) where the focus is on biomass valorisation. In both projects, SINTEF handled the process design starting from lab-scale data (from batch to continuous scale) and it carried out sub-models' generation, integration into a single flowsheet, up-scale, and optimization.

#### Acknowledgements

The authors are grateful to VALUABLE (grant agreement No 101059786) and Bio4Fuel (grant No 257622).

#### References

- Bio4Fuels, <https://www.nmbu.no/en/research/projects/bio4fuels> (accessed October 7, 2023).
- Bisotti F., Gilardi M., Berglihn O.T., Tschentscher R., Eijnsink V.G.H., Várnai A., Wittgens B., 2023, Soft modelling of spruce conversion into bio-oil through pyrolysis – Note II: pyrolysis, In: A.C. Kokossis, M.C. Georgiadis, E. Pistikopoulos (Eds.), *Computer Aided Chemical Engineering*, Elsevier, 2023: pp. 769–774.
- Caballero J.A., Grossmann I.E., 2008, An algorithm for the use of surrogate models in modular flowsheet optimization, *AIChE Journal*, 54, 2633–2650.
- Di Pretoro A., D'Iglio F., Manenti F., 2021, Optimal cleaning cycle scheduling under uncertain conditions: A flexibility analysis on heat exchanger fouling, *Processes*, 9, issue 1, 1-18.
- Dente M., Bozzano G., Faravelli T., Marongiu A., Pierucci S., Ranzi E., 2007, Kinetic Modelling of Pyrolysis Processes in Gas and Condensed Phase, In: G.B. Marin (Ed.), *Advances in Chemical Engineering*, Academic Press, pp. 51–166.
- Dussan K., Dooley S., Monaghan R.F.D., 2019. A model of the chemical decomposition and pyrolysis kinetics of lignin. *Proceedings of the Combustion Institute*, 37(3), 2697-2704.
- Galeazzi A., de Fusco F., Prifti K., Gallo F., Biegler L., Manenti F., 2024, Predicting the performance of an industrial furnace using Gaussian process and linear regression: a comparison, *Computer and Chemical Engineering*, 181, 108513.
- Gilardi M., Bisotti F., Berglihn O.T., Tschentscher R., Eijnsink V.G.H., Várnai A., Wittgens B., 2023, Soft modelling of spruce conversion into bio-oil through pyrolysis – Note I: steam explosion and LPMO-activated enzymatic saccharification, In: A.C. Kokossis, M.C. Georgiadis, E. Pistikopoulos (Eds.), *Computer Aided Chemical Engineering*, Elsevier, 2023: pp. 757–762.
- Humbird D., Davis R., Tao L., Kinchin C., Hsu D., Aden A., Schoen P., Lukas J., Olthof B., Worley M., Sexton D., Dudgeon D., 2011, Process design and economics for biochemical conversion of lignocellulosic biomass to ethanol: Dilute-acid pretreatment and enzymatic hydrolysis of corn stover. NREL technical report NREL/TP-5100-47764.
- Hansen L.D., Østensen M., Arstad B., Tschentscher R., Eijnsink V.G.H., Horn S.J., Várnai A., 2022, 2-Naphthol Impregnation Prior to Steam Explosion Promotes LPMO-Assisted Enzymatic Saccharification of Spruce and Yields High-Purity Lignin. *ACS Sustainable Chemistry & Engineering*, 10, issue 16, 5233-5242.
- Henao C.A., Maravelias C.T., 2011, Surrogate-based superstructure optimization framework. *AIChE Journal*, 57, 1216–1232.
- Martín M., Taifouris M., Galán G., 2023, Lignocellulosic biorefineries: A multiscale approach for resource exploitation. *Bioresources Technology*, 385, 129397
- Martín M., Grossmann I.E., 2012, Simultaneous Optimization and Heat Integration for Biodiesel Production from Cooking Oil and Algae. *Industrial & Engineering Chemistry Research*, 51, 7998–8014.
- Quirino P.P.S., Amaral A.F., Manenti F., Pontes K.V., 2022, Mapping and optimization of an industrial steam methane reformer by the design of experiments (DOE), *Chemical Engineering Research and Design*, 184, 349-365.
- Ranzi E., Cuoci A., Faravelli T., Frassoldati A., Migliavacca G., Pierucci S., Sommariva S., 2008. Chemical kinetics of biomass pyrolysis. *Energy & Fuel*, 22(6), 4292-4300.
- VALUABLE, <https://valuable-project.eu/> (accessed October 7, 2023)
- Velidandi A., Gandam P.K., Chinta M.L., Konakanchi S., Bhavanam A.R., Baadhe R.R., Sharma M., Gaffey J., Nguyen Q.D., Gupta V.K., 2023, State-of-the-art and future directions of machine learning for biomass characterization and for sustainable biorefinery, *Journal of Energy Chemistry*, 81, 42–63.