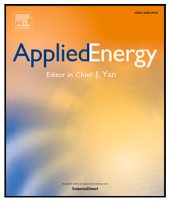
Applied Energy 306 (2022) 117954

Contents lists available at ScienceDirect

Applied Energy

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



Overcoming decision paralysis—A digital twin for decision making in energy system design

Julia Granacher *, Tuong-Van Nguyen, Rafael Castro-Amoedo, François Maréchal

Industrial Process and Energy Systems Engineering (IPESE), École Polytechnique Fédérale de Lausanne, Switzerland

ARTICLE INFO

Keywords:
Decision support
Multi-criteria decision analysis
Multi-objective optimization
Process and energy system design
Interactive optimization
Digital twin

ABSTRACT

The design of efficient energy systems, through the development of new technologies and the improvement of current ones, requires the use of rigorous process synthesis methods for generating and analysing design alternatives. We introduce a digital twin of process and energy system design that interactively translates needs and preferences of decision makers into an optimization-based model and generates meaningful solutions. The Interactive Digital Twin (InDiT) assists decision makers in steering the exploration of the solution space and guiding them towards relevant system design decisions, taking into account multiple aspects such as the impact of uncertainties and multi-criteria analysis. InDiT enhances step-by-step communication with the decision maker, relying on visual aids to keep the communication during solution generation and exploration intuitive and flexible. In this way, decision makers are guided towards relevant solutions and improve their understanding of relations between the problem definition and system design decisions, while InDiT builds on the decision makers' preferences and can, after training, suggest solutions that are best-suited to their interests. The novelty of this work lies in the holistic approach of addressing both (i) the systematic generation and exploration of solutions with the assistance of a digital consultant, which translates the decision maker's needs into machine language and vice versa, and (ii) the interactive step-by-step technique on filtering and evaluating solutions intuitively. This guarantees that the decision maker does not only get solutions based on the design specifications made, but that personal preferences are taken into account during the solution synthesis step, and that the solution space can easily be explored under different criteria. The proposed methodology is demonstrated and applied to the design case of an integrated multi-product biorefinery.

1. Introduction

1.1. Motivation and state of the art

Climate change mitigation is one of the most pressing challenges society is facing, revealing the need for efficient and reliable design methods of energy systems that are sustainable in economical, environmental and social terms.

Process systems engineering (PSE) aims to define, design, plan and control systems with physical, chemical or biological operations, such as energy systems or chemical processes [1].

Process simulation. As described by Stephanopoulos and Reklaitis [2], the first steps of PSE were dominated by process simulation and analysis, relying on complex flowsheeting methods. In recent years, surrogate models – which describe a complex system by a set of simplified but rather accurate equations – have been used as an alternative, especially in combination with complex and CPU-intensive optimization problems [3,4].

Process synthesis. The emerging focus on quantitative descriptions of processes and phenomena by means of simulation led to a more thorough analysis of the system performance and on the ways to improve it [5]. Following unit operations selection and interconnection definition, process synthesis is applied for further development, simulation and optimization. Two main types of models are developed: traditional sequential-conceptual models and systematic superstructure optimization-based ones [3].

Sequential models/hierarchical decomposition. Sequential models/hierarchical decomposition as proposed by Douglas [6] assume a natural hierarchy among engineering decisions made during the generation of a chemical process flowsheet [3]. Such approaches have the potential

* Corresponding author.
E-mail address: julia.granacher@epfl.ch (J. Granacher).

https://doi.org/10.1016/j.apenergy.2021.117954
Received 27 May 2021; Received in revised form 22 September 2021; Accepted 26 September 2021
Available online 8 November 2021

Nomenclature

Abbreviations

CAPEX	Capital expenditure
DM	Decision maker
DME	Dimethyl ether
FT	Fischer Tropsch
GWP	Global warming potential
IO	Interactive optimization
KPI	Key performance indicator
LHS	Latin Hypercube Sampling
LHV	Lower heating value
MADA	Multi-attribute decision analysis
MeOH	Methanol
MILP	Mixed Integer Linear Programming
OPEX	Operating expenditure
PSE	Process system engineering
TOPSIS	Technique for Order of Preference by Similarity to Ideal Solution
TOTEX	Total expenditure

Indices and sets

$\Theta \in \mathbf{D}$	Decision variables of upper level framework in InDiT, defining bounds of lower level framework decision variables
$j \in \mathbf{K}$	Number of objectives considered in optimization
$l \in \mathbf{L}$	Generated solutions
$m \in \mathbf{M}$	Number of points on Pareto-front
$n \in \mathbf{N}$	Formulations of the multi-objective optimization problem used for solution generation with InDiT
$p_u \in \mathbf{P}$	Parameters subjected to uncertainty
$t \in \mathbf{T}$	System states
$u \in \mathbf{U}$	Units considered in the superstructure

Parameters and variables

$c_u^{\mathrm{op1/2}}$	Specific fixed/variable operating cost of unit u, calculated for reference size $f^{\mathrm{mult}} = 1$, [USD]
C_u^{inv}	Investment cost of unit u, [USD]
C	Amount of carbon, [kg]
E	Amount of electricity, [kWh]
η_{en}	Energy efficiency, [–]
η_{carbon}	Carbon efficiency, [–]
f_u^{mult}	Unit sizing factor, [–]
$F_u^{\mathrm{min/max}}$	Lower/Upper bound of unit sizing factor, [–]
$Impact$	Global warming potential after IPCC2013, [kgCO_2eq/year]
n/i	Investment lifetime/Investment interest rate
p_{payback}	Payback period, [years]
t_t^{op}	Operating time per time step t, [s]
y_u^{use}	Binary decision variable on unit installation , [–]
$y_{u,t}^{\mathrm{use}}$	Binary decision variable on unit usage at timestep t, [–]
Y_u	Bound on unit installation, indicates if unit is considered, [–]
z	Objective functions considered in optimization

to drastically reduce the complexity of the synthesis problem, but make decision interactions difficult to integrate. Examples are the design of particular sub-systems such as distillation-based separation systems, or heat and mass exchanger networks [7,8]. Grossmann [9] added the aspect of system optimality and introduced a way of formally quantifying a system's configuration performance, which led to the second category of models: superstructure optimization-based approaches.

Superstructure optimization. Superstructure optimization consists of three steps: *(i)* definition of a network of all potential unit operations and connections, including the set of all feasible alternative process configurations, *(ii)* translation into a mathematical programming model and *(iii)* result generation by solving an optimization problem [1]. Since this paper focuses on the solution generation and exploration, the first is only mentioned here for completeness. The interested reader may consult [1] for a detailed overview of developed superstructure representations. Mixed Integer Programming is widely applied to energy and process system optimization, and the mathematical problem is formulated either as a linear or a nonlinear problem [9], using discrete and continuous decision variables, and one or multiple objectives. Typically, discrete variables describe structural alternatives, whereas continuous variables represent process conditions and equipment sizes. To ensure useful results, the generation of a set of feasible alternatives may be preferred to the generation of a single system design, which may be optimal only under certain external conditions [10].

Multi-criteria decision analysis. The approach of multi-criteria decision analysis is particularly relevant when one wants to analyze the trade-off between conflicting objectives, for example through multi-objective optimization, coupled with Pareto visualizations. Several methods exist for creating a set of alternative solutions: Multi-objective optimization is widely applied for analysing trade-offs between two or more objective functions. Applications of the method range from the design of biomass conversion and biorefinery applications [11,12] to the optimization of mass and energy networks in industrial processes [13,14] and numerous applications in urban systems [15,16]. The procedure for obtaining solutions from a multi-objective optimization formulation can be either of deterministic or of heuristic nature [17]. In heuristic-based methods, stochastic exploration of solutions is employed. The implementation is straightforward, but many iterations may be needed to reach what is often only an approximation of the Pareto-front [18]. Deterministic exploration guarantees Pareto-optimal solutions. In that context, the multi-objective problem is converted into multiple single objective functions by the application of parametrized scalarization functions, enabling the use of common single-objective solvers to explore the Pareto curve [19].

Another rigorous strategy to obtain alternative solutions is to calculate near-optimal solutions using integer cuts [20]. This method produces different solutions for a single-objective problem by continuously adding constraints on the integer variables, forbidding combinations of binaries that were previously chosen [21]. The concept of near-optimal solutions is particularly applied for the analysis of parameter uncertainty [22], for long-term energy system planning [23] and for characterizing properties of preferred solutions and correlations between optimizer decisions [24]. A general overview of optimization strategies for obtaining solutions is provided by Liu et al. [25]. Once a set of solutions is generated, each one can be characterized by a series of performance indicators. Choosing among proposed solutions requires multi-criteria decision analysis methods for screening solutions and identifying the most attractive ones. A comprehensive review of multi-criteria decision making methods is provided by Cajot et al. [26].

Interactive optimization. In recent years, interactive optimization (IO) has gained interest of the research community, impelled by the incorporation of expert knowledge and experience in the optimization process and solution exploration process, as well as intuition and personal preferences [17,18]. As a consequence, computational efforts can be significantly reduced, given that only the relevant space interesting to the decision maker is explored [27]. Besides increasing the decision maker's confidence in the obtained results, the interaction process

supports intuitive learning, without relying on a priori preference information [18]. However, IO methods rely on the availability of a human decision maker, their ability to devote time to the solution exploration, and their capability to understand the demands of the IO [18,19]. Thus, IO applications should be intuitive and flexible in adapting to the user's inputs [18].

Visualization is recognized as an important aspect of IO methods [17,18]. Parallel coordinates gained increasing attention in this regard, as they allow for intuitive comparison of scenarios [28]. Abi Akle et al. [29] approved their effectiveness in displaying multi-objective optimization results compared to other visual aids for exploring solutions, while Cajot et al. [18] applied them to multi-criteria decision support for urban energy systems. Reviews of interactive optimization and choices of visualization tools are provided by Branke et al. [19], Miettinen [30].

Uncertainty consideration. With the aim of generating meaningful solutions, the question of uncertainty in data availability certainly comes to mind. Assumptions on which investment decisions for industrial energy system are based, such as the development of equipment and resource cost, are often proven to be wrong retrospectively [31]. For providing assistance to informed decision making, it is crucial to take the uncertainty of these assumptions into account. Multiple studies were conducted to assess the impact of uncertainties in energy system designs. The review by Soroudi and Amraee [32] presents the state of the art in decision making under uncertainty applied to energy systems; a classification of different uncertainty handling methods is given, together with prospective advantages and disadvantages. Zhou et al. [33] present a review of decision analysis methods handling uncertainty. Uncertainty analysis is often embedded in the solution generation procedure via stochastic programming, generating solutions via a scenario-based approach. The encountered problems include the use of probability distribution functions that need to be available beforehand, as well as the problem size which quickly escalates and becomes untraceable [34]. As an alternative, robust optimization emerged to deal with the issue of large amounts of uncertain parameters and scarcity of data, considering the worst-case scenario of the uncertainty in the optimization problem [35].

1.2. Identified gaps and research objectives

In our research, we address the efficient generation of meaningful solutions for a decision maker by applying a digital twin concept to the solution synthesis and exploration for energy system and process design. In this context, *meaningful* refers to the preferences the decision maker expresses at any stage of the solution generation and synthesis.

Digital twins were originally proposed in the early 2000 and have been mainly applied as an enabling technology for smart manufacturing in industry for areas such as product design, production prognostics or health management [36]. Definitions of the concept found in literature range from virtual representations of objects or systems that interact with the physical counterpart throughout its life cycle to the seamless integration between cyber and physical systems based on simulation, real-time data, machine learning and reasoning [36,37]. Comprehensive reviews of digital twin definitions and applications are provided in [36,38]. In our research, we transfer the concept of a digital twin onto the decision making process an engineer passes through when designing a process and energy system. With the proposed **In**teractive **Di**gital **T**win (InDiT) methodology, we aim to guide a decision maker (DM) in multi-dimensional solution spaces based on user-defined preferences, providing insights about typical unit choices and sizes, correlations between user preferences and the solution space, as well as trade-offs between competing decision criteria. Other than the majority of computer-aided decision making approaches, our digital twin does not only synthesize results based on the technical system incorporated in a simulation or mathematical programming formulation, it moreover mirrors the whole decision process an engineer is passing through, translating their needs into computer language and retrieving relevant information.

Taking the DM's preferences into account at different stages allows for efficient solution generation, for adapting the search space if preferences change, and for exploring the solution space under different perspectives. The DM can explore solutions under varying criteria of interest and can adapt the search space of the optimization for solution generation accordingly. By coupling the visually-assisted solution exploration with a rigorous solution generation step relying on mathematical programming and optimization, we ensure generated solutions to be Pareto-optimal in a certain decision space. Tackling uncertainty in economic parameters at various steps of the solution generation and exploration, we allow for the analysis of solution performance under different economic conditions.

Our aim is not to provide the DM with one optimal solution, but to rather suggest a methodology that enables a fast and comprehensive exploration of the solution space of a complex system engineering problem, like in our case an integrated biorefinery pulp mill superstructure. The results proposed by InDiT mirror the DM's preferences and allow to analyze under which external criteria the generated decisions hold. To the knowledge of the authors, there is no comprehensive methodology that includes an interactive and adaptable solution generation exploration, taking DM input such as uncertainty distribution at different stages into account.

2. Methodology

2.1. Problem statement and overview of the proposed methodology

Our **In**teractive **Di**gital **T**win InDiT assists the generation and exploration of meaningful solutions for a given process synthesis superstructure, letting the DM influence the decision making at different stages of the process by translating their needs into computer language and steering the solution synthesis in the required direction. InDiT's worklflow consists of three main parts:

- Superstructure formulation
- Solution synthesis
  - Decision space exploration
  - Problem formulation
  - Solution generation
- Solution exploration and steering of result generation

The simplified workflow of the InDiT algorithm is presented in Fig. 1. The first step, namely the superstructure formulation, considers the system design as a list of units that need to be selected, sized and connected to realize a system functionality under user-predefined conditions and objectives. The developed superstructure contains all considered units and possible connections between them. Based on steering inputs specified by the DM (shown as green boxes in Fig. 1), InDiT explores the decision space and defines a set of problems to be solved. Systematic solution generation strategies are followed by a comprehensive exploration and assessment of the generated results with regards to the system performance indicators. The solution exploration identifies solutions that are most promising and meaningful to the DM, and proposes steering actions on how to enrich the list of relevant solutions.

2.2. Superstructure formulation

The process superstructure describes the system's units and the way they can interact with others. By activating certain units and their connections, different system configurations are generated. For using the superstructure for the generation of solutions and evaluating

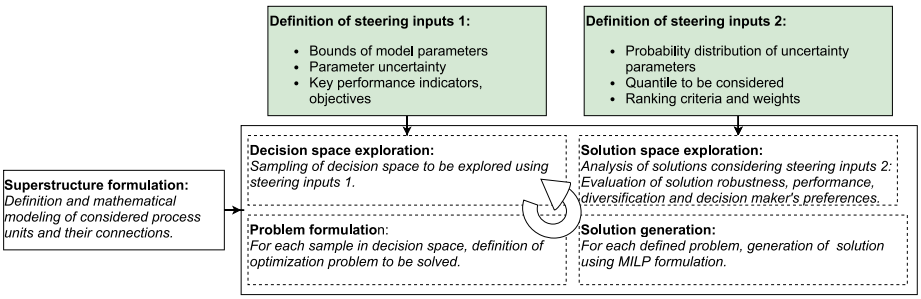


Fig. 1. Simplified flowchart of proposed methodology InDiT for the synthesis and exploration of solutions for superstructure optimization.

them based on economical, thermodynamic and environmental performance, the model needs to be developed in a systematic manner. The superstructure model encompasses two dimensions: the mass and energy conversion in the units and their integration in the system. The methodology for superstructure modeling and optimization is adapted from Gassner and Maréchal [11] and Kantor et al. [39]. From now on, it is referred to as the lower level framework in a upper/lower level structure, where the lower level framework generates results for optimization problems and is controlled by the upper level framework InDiT, where optimization problems are imposed, the results of the salve framework are evaluated and the generation of new results is steered.

The synthesis of configurations from the superstructure includes the manipulation of the decision variables by an optimizer. Decision variables of the lower level framework are of two types: they are either binary (installation/usage of a given unit y_u^{use}, $y_{u,t}^{\text{use}}$) or continuous (size/capacity factor of unit f_u^{mult}, $f_{u,t}^{\text{mult}}$). For a problem communicated by the upper level framework InDiT, the lower level framework generates a solution and reports it back to InDiT. In the lower level framework, the decision variables for a solution are fixed solving a mixed-integer linear programming (MILP) problem formulated in the AMPL optimization language [40], using the CPLEX branch-and-bound algorithm [41]. Fig. 2 shows the simplified structure of the lower level framework, including the superstructure description and the solution evaluation for each problem formulated by InDiT. Further information on the mathematical formulation of the superstructure model applied in the lower level framework can be found in Appendix.

2.3. Solution synthesis

The solution synthesis of InDiT consists of the definition of steering inputs, the exploration of the decision space, the problem formulation and the generation of solutions using the lower level framework. It is followed by the exploration of the solutions (Fig. 3).

2.3.1. Definition of steering inputs

The aim of defining steering inputs for InDiT is to capture the DM's preferences and needs, and to translate them into framework conditions useable for solution synthesis.

Bounds of the decision variables and uncertainty in parameters. Firstly, D relevant variables of the upper level framework that influence the decision space of the lower level framework are identified, containing the bounds of the continuous (x) and discrete (y) decision variables of the lower level framework. The discrete decision variables of the upper level framework Θ_y set the upper bound $Y_u \geq y_u^{\text{use}}$ of the installation of independent units in the lower level framework and therefore determine whether a unit can be considered for solution generation. The continuous decision variables Θ_x contain the upper bounds of the unit multiplication factor $F_u^{\max} \geq f_u^{\text{mult}}$ for independent units. Independent in that sense implies that the unit size is not directly correlated to the size or operation of another unit. Furthermore, lower level framework parameters p_u that are subjected to uncertainty, such as the equipment lifetime or the electricity price, are identified with their ranges of variation.

Objectives and key performance indicators:. Objectives z that should be evaluated by the lower level framework are identified, as well as user-defined key performance indicators. The difference between objectives and key performance indicators (KPIs) is that objectives are used to steer the lower level framework optimization (e.g. minimizing the total system cost), whereas a KPI is an indicator of special interest for a DM (e.g. share of energy from renewables). As we are using an MILP formulation in our lower level framework, the objectives have to be of linear nature, while KPIs can include nonlinear relations as well, since they are computed a posteriori. Objectives and KPIs can be competing with one another.

2.3.2. Decision space exploration

The aim of the exploration of the decision space is to enhance the generation of a variety of valid solutions by the lower level framework. For this purpose, samples are drawn from the identified upper level decision variables Θ. Furthermore, the parameters p_u for which an uncertainty range has been identified are sampled within an identified distribution. Latin Hypercube Sampling (LHS) is applied. This stratified sampling technique splits the range of each input variable into intervals of equal probability, and each of the intervals is sampled once [43]. The sampling of the decision variable bounds and of the uncertain parameters leads to a set of N formulations of a multi-objective optimization problem, consisting of P uncertain parameters and D decision variable bounds of the lower level framework.

2.3.3. Problem formulation

The problem definition step corresponds to the translation of the decision space characteristics, containing objectives, decision variable bounds and uncertainty parameters into optimization problems solved by the lower level framework. The MILP problem on the lower level framework that is characterized by the inputs of the upper level framework for each problem can be summarized as follows:

$$\begin{aligned} &\min\ z(x,y,p) \\ &\text{subject to} \\ &\quad g(x,y,p) \leq 0 \\ &\quad h(x,y,p) = 0 \\ &\quad \Theta_{x,\min} \leq x \leq \Theta_{x,\max}, \Theta_x \in R^D \\ &\quad \Theta_{y,\min} \leq y \leq \Theta_{y,\max}, \Theta_y \in R^D. \end{aligned} \tag{1}$$

$z(x,y,p) \in R^K$ contains the K objective functions that are to minimize, x, y are the continuous and discrete decision variables of the lower level framework, p the parameters, $g(x,y)$ define the inequality constraints, and $h(x,y)$ the equality constraints, while Θ defines the bounds of the decision variables imposed by InDiT.

Optimization approach. We choose to evaluate our defined problem for each sample by applying a deterministic optimization procedure, as it guarantees Pareto-optimal solutions with fewer iterations than with heuristic methods. This requires the transformation of the multi-objective problem formulation of Eq. (1) into parametrized single objectives.

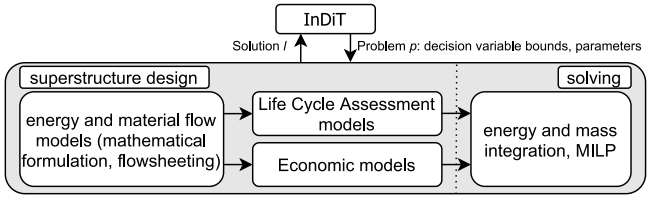


Fig. 2. Lower level optimization approach, integrated in InDiT, adapted from [11,42].

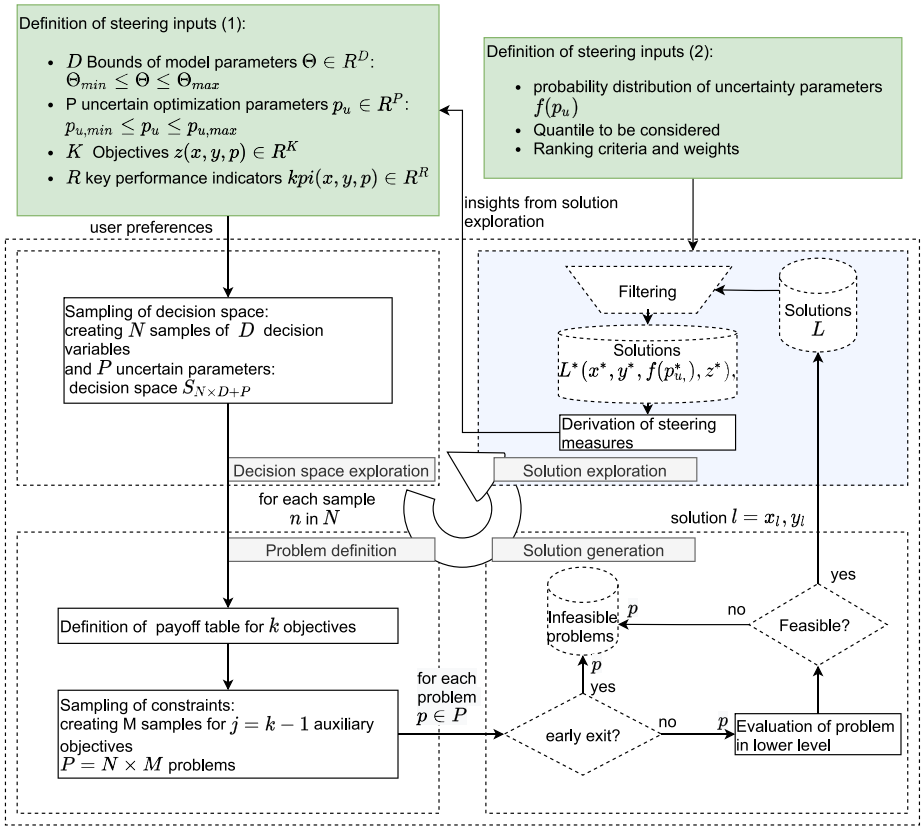


Fig. 3. Proposed methodology InDiT for the synthesis and exploration of solutions for superstructure optimization. The solution exploration is specified further in Fig. 5.

Scalarization function. For transforming the problem into a single-objective formulation, we apply an adapted form of the ε-constraint method, first introduced by Haimes et al. [44] and adapted by Mavrotas [45] and Cajot et al. [18].

Compared to other common methods such as the weighted sum method, the ε-constraint method is able to handle convex as well as non-convex Pareto-fronts, does not suffer from counter-intuitive weight specification and the variation of the parametrized constraints leads to a richer and more diverse set of solutions [18,19,45,46].

When applying the ε-constraint method, only one objective k is minimized and the others j are transformed into inequality constraints characterized by an upper bound ε:

$$\begin{aligned} &\min_{x,y} \quad z_k(x,y,p) \\ &\text{subject to} \\ &\qquad z_j(x,y,p) \leq \varepsilon_{m,j}, j = 1,\ldots,K-1, j \neq k, \\ &\qquad g(x,y,p) \leq 0 \\ &\qquad h(x,y,p) = 0 \end{aligned} \tag{2}$$

where $k \in 1....K, m = 1....M$, M being the total number of points on the Pareto-front.

Payoff table calculation. The range of each parametrized objective function in which the upper bounds ε can be varied needs to be identified when applying the ε-constraint method. In the original ε-constraint approach, this is done by calculating the individual optima of each objective function. However, this may lead to weak Pareto-optimal points in case of multiple alternative solutions [45]. In our approach, the payoff table between the different identified objectives and thus the ranges of the ε-constraint are calculated applying an augmented ε-constraint method as suggested by Mavrotas [45], considering lexicographic optimization. The objective with highest priority to the DM is optimized obtaining $min(z_1) = z_1^*$. The second function is optimized after adding the constraint $z_1 \leq z_1^*$, and $min(z_2) = z_2^*$ is obtained. Both constraints $z_1 = z_1^*$, and $z_2 = z_2^*$ are added as constraints for optimizing z_3, until all objective functions are treated. This guarantees that all the payoff table is built with only Pareto-optimal solutions, therefore excluding weak Pareto-optimal ones (Fig. 4B, see dominated ε bounds).

Sampling of ε-constraints:. After the non-dominated payoff table is calculated, a sampling algorithm is applied to define M bounds for the parametrized objective functions z_j.

Burhenne et al. [47] compared different sampling algorithms and concluded that the Sobol approach [48] provides a robust and efficient exploitation of the search space. The approach has also been explored by Copado-Méndez et al. [49] and Cajot et al. [18]. Sobol sequence sampling compared to systematic sampling is illustrated in

Fig. 4A, where the exploration of the search space in a $k-$dimensional optimization problem is shown, proving the superiority of the former. Compared to systematic sampling, it obtains a faster representation of the Pareto-front.

In this work, as our aim is to allow for flexible and efficient exploration of the solution space for multiple objective functions, we apply Sobol sampling to create parametrized constraints $\varepsilon_{m,j}$ for generating solutions.

2.3.4. Solution generation

For each of the $N \times M$ samples of Θ, p_u and ε, the model is evaluated and solutions are generated in a deterministic way using the lower level framework. We ensure that generated solutions are better than a benchmark set of solutions generated for the least-favourable set of economic parameters in p_u under relaxed conditions for the decision bounds Θ. Thus, the generated benchmark solutions are the optimal ones for the worst possible economic scenario. For this purpose, we determine the least-favourable set of economic condition when trying to minimize total cost. This means that if the correlation between total cost and the parameter $p_{u,i}$ is 1, the parameter is set to its maximum, and if the correlation is -1, it is set to its minimum in the distribution sampled before. For example, increasing the cost of natural gas increases the total cost, so the former is set to its maximum. For selling price of a product, the correlation is reversed, so all cost for sold products are set to their minimum. For the resulting set of least-favourable economic parameters, we calculate the range of the ε-constraints. The decision maker can define a tolerance range with which obtained results can exceed the calculated ranges in the economic worst case, relaxed scenario. For each sample n, InDiT will check if the calculated payoff table is fulfilling the criteria of being in the defined tolerance. This means that when minimizing objective j for the payoff table calculation as a single-objective optimization problem, the result needs to be beyond a defined threshold for the same objective obtained in the worst case relaxed scenario. If for a sample n any of the single objective optimizations in the payoff table definition are accepted, the sample is considered for further solution generation. If all single objective optimizations perform worse than the defined relaxed economic worst case, the sample is discarded (demonstrated for two objectives in Fig. 4B).

In order to avoid the evaluation of infeasible sets of ε, an early exit mechanism is implemented in line with what was suggested by Mavrotas [45]. This is notably important when more than one objective is subjected to an ε-constraint, as some propositions of constraints might be infeasible. Assuming all objectives are to be minimized, if the evaluation of the m^{th} ε sample is infeasible, it is added to a set of infeasible samples. For each new sample, the closest infeasible sample is chosen based on the distance between the ε of the first constrained objective and the respective ε in infeasible samples. Then, if any of the other εs of the sample is smaller than the respective epsilon of the infeasible sample, the evaluation is discarded. For example, let our optimization contain three objective functions, leading to an ε-constraint formulation of parametrized functions $z(x,y)_1 \leq \varepsilon_1$ and $z(x,y)_2 \leq \varepsilon_2$. If now our closest set of infeasible ε-constraints is $\varepsilon_{infeasible} = [\varepsilon_1^*, \varepsilon_2^*]$, for which we know that $\varepsilon_1^* \geq \varepsilon_1$ and if $\varepsilon_2^* \geq \varepsilon_2$, the sample must be infeasible as well and is thus not evaluated, but instead added to the set of infeasible samples.

2.4. Solution exploration

The described algorithm leads to a set of L solutions, each being part of a Pareto-front for a set n of model parameters. For analysing the solutions and enable informed decision making, methods for systematically ranking and visualizing the solutions based on desired criteria are coupled with DM input that steers the decision support in the desired direction. The proposed workflow for exploring solutions with InDiT is illustrated in Fig. 5. On the left side, the flow of the algorithm is illustrated, while on the right side the individual steps included are summarized. Trapezoidal boxes indicate a filtering step in which the selected solutions are reduced by some criteria, while squared boxes represent data processing performed for further analysis. All steps included in the exploration aim at guiding the DM towards a relevant subset of solutions. The steps can be repeated for different steering criteria and DM inputs for comparing results.

The main concepts of each step are presented in the following paragraphs. Details on the implementation can be found in Appendix.

2.4.1. Filtering solutions based on decision maker's preference

One way to analyze solutions is to visualize trade-offs of the desired objectives and KPIs in Pareto-fronts. This is particularly interesting when one wants to balance solutions based on competing objectives, such as for example investment cost and expenses for operation. Furthermore, the solutions can be ranked based on objectives and KPIs, as it was done by Celebi et al. [50]. In our approach, we enable the evaluation of solutions based on a variety of KPIs and objectives that the DM might be interested in. As the consideration of multiple criteria can result in a complex and nontransparent set of data that is difficult to display in table format, we chose graphical representation to allow for a first overview of the solution space. In a first filtering, unique solutions are identified by grouping all solutions L on optimizer decisions. For each unique solution, the occurrence is calculated based on how often the optimizer chooses the respective configuration. All unique solutions are displayed in parallel coordinates, which allows for comprehensive analysis of the trade-offs between different process pathways. In parallel coordinates, negative correlations are displayed via crossing lines, while synergies are indicated by non-crossing lines [18,51,52]. Colour indication and other visual encoding such as line-style can be used to display characteristics of the data. In this first graphical representation, we show the user-defined KPIs as well as the objectives in one single plot. At this point, the DM has the option to refine the selection of solutions by manipulating the visualization tool. Brushing an axes leads to the display of the solutions within the desired range [53,54]. Dismissing axes enables the DM to solely focus on KPIs of interest, preventing unnecessary distraction information [18]. The DM also has the option to display the underlying optimizer decisions in the parallel coordinates. After the first user-defined filtering of solutions, the solution space L is reduced to $L^* \leq L$.

2.4.2. Evaluation of operating robustness of solutions

After reducing the solution space based on DM preferences, the robustness under varying economic conditions of the remaining solutions is analyzed. Several studies [24,31,55] have noted how modeling work is subjected to uncertainty in the input assumptions, and that bad premises will lead to bad results, no matter how well-defined a model is. The previously calculated occurrence of a solution in the solution space is only a proxy of performance under different economic scenarios, provided solution generation also takes different ε-constraints and decision spaces into account. To prevent misjudgements from results which are weak as in their performance does not hold under variations in economic parameters, a detailed uncertainty analysis is included, which can be completely steered by the DM. In the first steering input, the DM has already defined the parameters deemed uncertain (p_u), as well as the ranges in which they vary in order to generate a diverse set of solutions. For this purpose, near-random distribution of parameters is assumed per default in InDiT, ensuring even exploration of the decision space. In the second steering input used for solution exploration under uncertainty, probability distributions functions $f(p_u)$ need to be defined. For a given distribution of p_u, InDiT re-calculates the resulting distribution of the economic KPIs per solution. This distribution can be used to derive insights on how well a solution performs under different economic scenarios. For quantifying this performance, we enable the visualization of the obtained distribution with a user-defined percentile and the mean of the obtained distributions. Per default, we include the 95% percentile of the recalculated KPIs to be

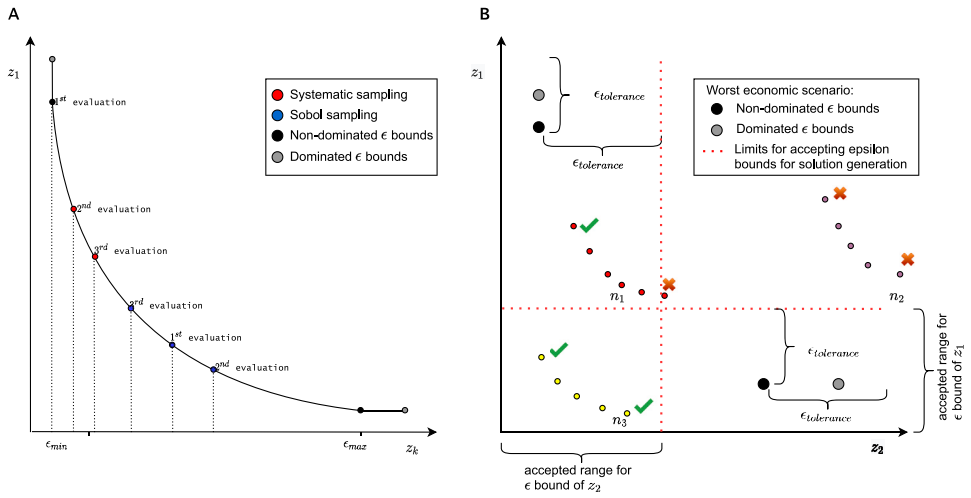


Fig. 4. **A:** Systematic vs Sobol sampling of Pareto-front. Comparison of systematic (red) to Sobol (blue) sampling approach for performing ϵ- constraint method minimizing 2 objectives. $\epsilon_{min/max}$ are the bounds of the parametrized function z_k, z_1 is the main objective, adapted from [18] **B:** Acceptance criteria for payoff table calculation. Non-dominated, worst economic scenario payoff table including decision maker's tolerance illustrated in black, dominated payoff table in grey. Sample n_1, n_3 are accepted and evaluated using epsilon bounds, n_2 is discarded after calculation of payoff table.

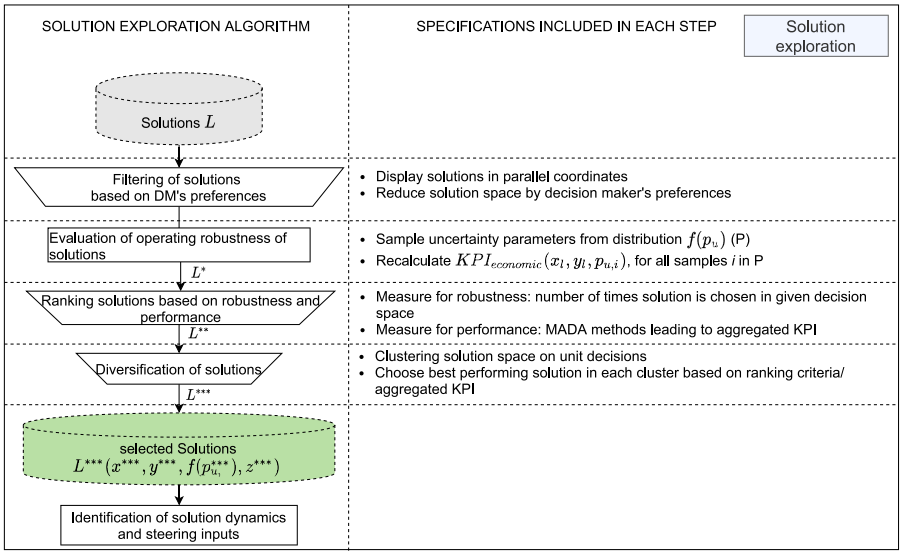


Fig. 5. Solution exploration in InDiT.

displayed in the parallel coordinates. Another visualization aiding-tool is the display of economic trade-offs in a Pareto-front, including the calculated distributions and the original data. That way, the DM can identify solutions that are Pareto-optimal for a desired percentile of the calculated distribution, and can also judge the impact of uncertainty on the economic KPIs. It is also possible to perform multiple analyses with the generated solutions, and compare the preferred ones with and without uncertainty.

2.4.3. Ranking solutions based on robustness and performance

Even though parallel coordinates facilitate the analysis of the solution space compared to tables, the multitude of KPIs and objectives to be considered still leads to a rich and complex decision space which is hard to capture in parallel coordinates. For this reason, we suggest to include a ranking mechanism for the performance of a solution a DM might be interested in. Solutions can be ranked based on a multitude of methods. For example, for each KPI of interest, solutions can be ranked on their best value, and a global ranking can be computed based on a desired weight of each of the relevant KPIs and objectives. Such aggregative multi-attribute decision analysis (MADA) methods are widely used in research and industry to rank solutions based on a defined metric of performance [18]. In our MADA approach, each unique solution is attributed with a score that supposed to draw the decision maker's focus on a limited number of information, not to replace all information gained form the rich display in parallel coordinates [18]. According to [18], most MADA approaches are based on the assumption that evaluation criteria can compensate each other. Therefore, we apply the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) method, which ranks each solution based on its proximity to an ideal solution and a worst solution. The ideal solution is characterized by the best value in every criterion, and the worst solution by the worst values of all criteria [56,57]. The TOPSIS method results in a ranking of solutions between 0 and 1, which can be added as an axis to the parallel coordinates plot. For applying TOPSIS, the DM needs to define the ranking criteria and associated weights as well as information about benefit or cost of each criteria to qualify the solutions with. These criteria can include not only the objectives and the KPIs, but also other solution properties, such as the size of a certain unit, the level of autonomy of the system, or the attributed occurrence. The occurrence of a unique solution can be seen

as a relative measure for how stable the respective set of decisions is, namely, how often it is selected for different decision spaces and economic scenarios. One key aspect of TOPSIS is that the evaluation criteria and their respective weights can be defined after the first set of solutions is generated, and distinct scores can be calculated based on different criteria that one might consider as important. After getting a better overview of the quality of the generated solutions with regard to the selected evaluation criteria, the DM can safely exclude some solutions that are uninteresting in their current purpose and receive a solution set $L^{**} \leq L^{*} \leq L$.

2.4.4. Diversification of solutions

Once all remaining solutions are ranked based on user-defined criteria, the question remains on how to chose the most promising ones. Apart from user-specific ranking criteria and performance under uncertainty, the richness of the solution space can be interesting to the DM. Having a final set of solutions that perform identically regarding the ranking and robustness measures, but are diverse in optimizer decisions, allows to choose one final solution based on more "soft" criteria, such as the current market situation (purchasing prices) or other subjective demands. Clustering solutions in groups with similar characteristics is a common approach to select configurations from a large set of solutions [58,59]. The solutions in L^{**} are clustered on the optimizer decisions, so that clusters of similar solutions are generated. K-medoids was used for clustering; it leverages on its robustness and relies on the existing dataset for representative solutions [60], unlike similar algorithms such as k-means. In addition, despite the higher computational effort, the suggested representative solutions are necessarily feasible, which might not be the case for the virtual cluster centroids calculated with k-means. For determining the optimal number of clusters, the elbow method [61] is employed. Principal component analysis is used for translating the optimizer decisions and the resulting clusters into 2 dimensions that can easily be visualized. For each of the clusters, the best solution – taking the user-defined rank into account – is chosen and added to the pool of final solutions L^{***}. That way, the DM can be sure to have a diverse solution space that contains the preferred solutions considering their personal evaluation criteria (see Fig. 6).

2.4.5. Understanding solution dynamics and refining the search

The remaining question is whether the solution space is sufficiently investigated, or the decision space needs to re-sampled for creating additional relevant solutions. For this purpose, it is necessary to understand how the optimizer decisions are influenced by choices expressed in the steering inputs, and how the desired KPIs are influenced by the optimizer decisions. Questions to be addressed in this stage of the algorithm are: Is our solution space explored sufficiently, or might there be other solutions interesting for the DM that have not yet been created? What decisions are always taken in the selected solutions, regardless of the steering parameters? Which decisions influence the KPIs the most, especially the ones interesting to the DM? Where do we need to re-sample if we want to create more meaningful solutions?

So, more general, the questions asked here are addressing the correlations between KPIs and decisions characterizing a solution. Answering them can reveal insights not only regarding the present solutions and their typical characteristics, but also help understanding the dynamic of solution generation and how one can influence it for generating more attractive solutions. Furthermore, the analysis of correlations also enables to decide which unit decisions are not influential or, similarly, that are always chosen at a constant size, revealing non-binding constraints for a selected KPI (see Fig. 6).

For validating the completeness of the solutions space and exploring KPI-specific characteristics, we propose to analyze which constraints are active, e.g. which constraints trigger changes in KPIs. We calculate the correlation matrix for the user-defined aggregated KPI and the optimizer decisions. Continuous decisions (unit choices) with the highest correlation are displayed in a bar graph, where the observed range of the decisions in all generated solutions L is shown, as well as the range of decisions present in the selected subset of solutions, L^{***}. This enables the decision maker to limit the decision space only to the ranges of interest for the selected solutions. If it is found that the whole range between the upper and lower bound for a unit size is used, it is advised to increase the bounds and relax the constraint in the steering input, if physically possible. On the other hand, if only a certain range or only one point is used for the selected solutions, the bounds can be changed to narrow down the range the lower level framework has to explore for generating solutions.

After it has been decided that the solution space is explored sufficiently, the solutions L^{***} are displayed in parallel coordinates, so that the decision maker can make its final choice. By evaluating the performance of the different solutions regarding desired KPIs, the final decision might be evident. If it is not evident, the DM can take other subjective factors into account.

2.4.6. Summary of the proposed methodology on exploring solutions

The methodology proposed is not supposed to replace a human DM, but rather to suggest and objectively (quantitatively) support the decision making process. The steps in the exploration of solutions allow for the efficient analysis of different aspects of embedded solution characteristics, simultaneously using the generated information for filtering. In this regard, InDiT serves as a translator and consultant between the computer-generated results and the information the DM needs for the decision process. Even though InDiT might not lead to one clearly dominating solution, it ensures that the final set of solutions are all within the accepted range of what the DM seeks, so they can focus their decision purely on personal preference, without worrying about meaningful information that might be hidden in the data without being considered. The analysis is enhanced by the option of including economic uncertainty and by the fact that the DM has the option to evaluate the same solution pool under different criteria and to assess trade-offs.

3. Application and results

The proposed methodology is intended for decision support in complex, multi-dimensional superstructure optimization problems, where the direct prioritization of preferences before solution generation is challenging to realize. It is intended for setting a framework of acceptable solutions at the beginning, and then exploring the generated results under different criteria. In that regard, we want to show that the method can be used to explore results assuming different perspectives. In our case study, we apply the proposed methodology of solution generation and exploration to the design of an integrated biorefinery pulp mill for the combined production of heat and fuel. Integrated multi-product biorefineries have been modelled and analyzed extensively in literature applying the above-described concepts of process system synthesis and optimization [50,62–65], also in relation to pulp mills [66–72]. However, to the knowledge of the author, available work does not follow interactive optimization approaches involving the DM actively in the decision process, but rather aims at presenting solutions for a given superstructure and optimization objectives. Furthermore, no study was found in that domain where solution generation and exploration is combined, allowing for the exploration of a solution space under different perspectives.

3.1. Process superstructure development

Biofuels represent an attractive low-carbon alternative for energy demands in sectors that can hardly be electrified, such as aviation or freight transportation. The pulp and paper industry is one of the largest energy consumers of the European industrial sector [73]. Integrating biorefinery concepts in an existing mill enables the plant to improve

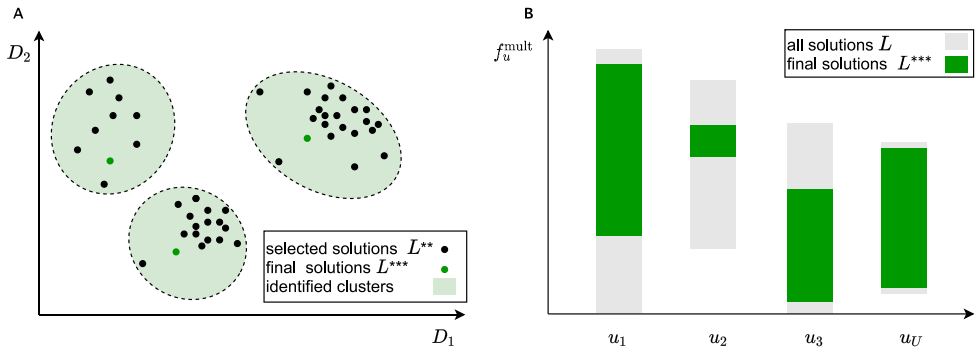


Fig. 6. **A:** Clusters based on decisions in selected solutions. Final solutions chosen in each cluster based on user KPI, displayed for decision space reduced to 2 dimensions. **B**: Bar plots of decisions (unit sizes) with highest correlations to user KPI. Grey: all solutions, Green, final solutions after clustering.

its resource and energy efficiency, while building a diverse and robust product portfolio of biofuels and pulp that can be provided to the market. Most of the European pulp and paper mills process wood as the basic raw material, which mainly consists of cellulose, hemicellulose and lignin [74]. In Fig. 7, a simplified block flow diagram of a generic Kraft process is presented. Wood is chopped, steamed and screened and fed into a cooking unit (digester), where cellulosic fibres are removed with the help of cooking chemicals [42]. The pulp exiting the digester is screened, bleached and dried. The weak black liquor that is exiting the digester is fed through a recovery unit consisting of an evaporator, a concentrator and a recovery boiler. The concentrator yields black liquor with a dry solid content of 75%, which is burnt in the recovery boiler. That way, the energy content of the burnt organic materials is recovered to run turbo-generators and to satisfy steam demands of the mill. Furthermore, inorganic substances are recovered as chemical pulping agents together with other by-products [42]. Inorganic residues form a smelt consisting of sodium carbonate (Na_2CO_3) and sodium sulphide (Na_2S), which is further used to produce green liquor. Green liquor is passed through a causticizing process where it is transformed to white liquor, consisting of sodium hydroxide ($NaOH$) and Na_2S. White liquor is then reused in the digesting process. Even though recovery boilers have proven over the last decades to be a relatively mature and reliable technology to process spent cooking liquors and produce steam and electricity for process use, the pulp and paper industry has been exploring alternatives [74]. The two main residue streams considered for fuel production are bark and black liquor. For the bark stream, dry gasification, with different fuel synthesis steps at the end is considered in the superstructure development. For adjusting the hydrogen to carbon monoxide (H_2/CO) ratio to the desired fuel synthesis, a water gas shift unit is included. The black liquor can – instead of being sent to the recovery boiler – be used in hydrothermal gasification for the generation of hydrogen. After gasification, different gas cleaning options, namely Pressure Swing Adsorption (PSA), physical absorption (Selexol) coupled with PSA and chemical absorption (Monoethanolamine, MEA) coupled with PSA are included. The purified hydrogen can then be used for adjusting the producer gas H_2/CO ratio. Three types of fuel can be generated, namely Fischer–Tropsch (FT) fuels, methanol (MeOH) and Dimethyl ether (DME). Optional units for increasing the hydrogen production and/ or de-bottlenecking the process are water or brine electrolysis yielding hydrogen, and co-electrolysis yielding producer gas that also needs adjustment regarding the H_2/CO ratio. When integrating biorefinery concepts, dark blue processes in Fig. 7 remain unchanged, as a constant production rate of pulp is assumed. However, the flows of black liquor and bark will be treated in alternative processes; consequently the energy requirements of the evaporator, concentrator, recovery and bark boilers are due to change. Furthermore, the load on the chemical recovery and the lime kiln is dependent on the amount of recovered cooking chemicals and will therefore be influenced, as the light blue units in Fig. 7 indicate. Thus, limitations to integrate biofuel production from bark and black liquor are the undisturbed operation of the pulp production, with no penalization on the performance, the supply of the heat demand of the pulp production, and the recovery of the cooking chemicals in an appropriate form.

The process models of the biorefinery units are built using Belsim Vali, a flowsheeting software [75] and integrated into the lower level framework, together with the pulp mill unit models. The reader is referred to Table 4 in the Appendix for more information on the modeling references and assumptions for the described superstructure.

3.2. Solution synthesis

3.2.1. Definition of steering inputs

Bounds of decision variables. Table 1 provide an overview of the decision variables inherent to each unit, where the unit multiplication factor f^{mult} and the upper bound $F^{\max}$ of the lower level framework optimization are given. If $F^{\max}$ is given as a range, it means that the maximum size of the particular unit is sampled in the steering input 1, so $F^{\max}$ is part of the D decision variables of the upper level framework, Θ, to generate a variety of results. The same applies for Y_u that decides the particular unit is considered (1) in the model or not (0). While some units are always considered, Y_u is varied in the steering input for others, in order to make the solution space more diverse. The lower bound of the unit sizes $F^{\min}$ is set to zero. Table 1 summarizes the decision variables of our case study. For generating a wide set of results, $F^{\max}$ ranges are varied for two of the independent process unit decisions in feasible domains for unit sizes.

Economic parameters subjected to uncertainty. For our case study of an integrated pulp mill biorefinery, we choose the economic parameters in Table 2 to be subjected to uncertainty. C_u^{inv} refers to the investment cost of a certain unit. For demonstrating the method, we assume variation of 10% from the reference cost for all equipment parameters. However, for future analyses, other user-defined variations might be applied, inspiration in this regard can come from [76,77]. Investment cost for all units are taken from nonlinear correlations provided by [76,78,79]. Operating cost are taken from various sources listed in Table 2. For the distribution of the operating cost, references in literature suggest different shapes such as normal or beta [55]. However, in order to avoid bias from strong assumed distributions, we chose to sample all distributions as random uniform in the ranges given in steering input 1, Table 2.

Key performance indicators. Economic and environmental criteria are chosen to evaluate system configurations. As described in Section 2.3, they are not intended to be used as objectives for generating solutions, but rather for comparing generated solutions based on competing criteria of interest for the decision maker.

OPEX: For the calculation of the operating costs, c_u^{op1} represents specific fixed annual operating costs and c_u^{op2} accounts for the specific

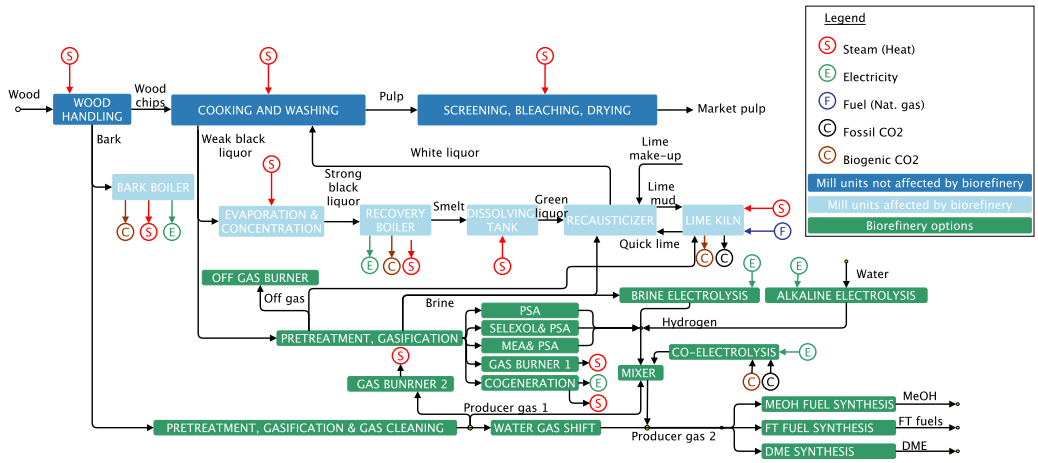


Fig. 7. Simplified block flow diagram of Kraft process with integrated biorefinery. Dark blue units are considered a constant size and operation, whereas light blue units change their load depending on the integration characteristics with the biorefinery (green).

Table 1
Relevant decision variables of the lower level framework optimization and steering inputs of the upper level framework. For each unit in the superstructure presented in Fig. 7, the multiplication factor f^{mult} with which the reference size is multiplied is given, as well as the upper bound $F^{\max}$ and the discrete parameter Y_u that decides whether a unit is considered in the solution generation. If ranges are given for $F^{\max}$ and Y_u, it means that these variables are included in the decision variables of the upper level framework Θ and are sampled for generating a rich solution space.

Unit	f^{mult}	$F^{\max}$	Reference size $f^{\text{mult}} = 1$	Y_u
Pretreatment, gasification gas cleaning	m_pretreatment	2	1 kg/s bark	1
Water gas shift	m_wgs	0–2	1 kg/s producer gas	1
MeOH fuel synthesis	m_meoh	5	1 kg/s producer gas	0/1
DME synthesis	m_dme	5	1 kg/s producer gas	0/1
FT fuel synthesis	m_ft_syn	5	1 kg/s producer gas	0/1
Pretreatment, gasification	m_chtg	20	1 kg/s black liquor	0/1
PSA	m_psa	7	1 kg/s producer gas	1
MEA & PSA	m_mea_psa	7	1 kg/s producer gas	1
Selexol & PSA	m_sel	7	1 kg/s producer gas	1
Mixer	m_mixer	2	1 kg/s producer gas	1
Brine electrolysis	m_br_elec	850	0.0005 kg/s hydrogen	0/1
Alkaline electrolysis	m_alk_elec	75	0.00534188 kg/s hydrogen	0/1
Co-electrolysis	m_co_elec	0.5	3 kg/s producer gas	0/1
Cogeneration	m_opx1	–	1 kg/s producer gas	0/1
Gas burner 1	m_burn1	–	1 kg/s producer gas	0/1
Gas burner 2	m_burn_pg	–	1 kg/s producer gas in	0/1
Lime Kiln	m_lime_kiln	–	1 kg/s quicklime out	1
Off gas burner lime kiln	m_burn_lk_og	–	1 kg/s offgas in	1
Natural gas burner lime kiln	m_burn_lk_gas	–	1 kg/s natural gas in	1
Recausticizing	m_recaust	1	2.16 quicklime in	1
Evaporator & concentrator	m_evap_conc	65	1 kg/s weak black liquor	1
Recovery boiler	m_boiler_rec	15	1 kg/s concentrated black liquor	1
Bark boiler	m_boiler_bark	0–2	1 kg/s bark	0/1

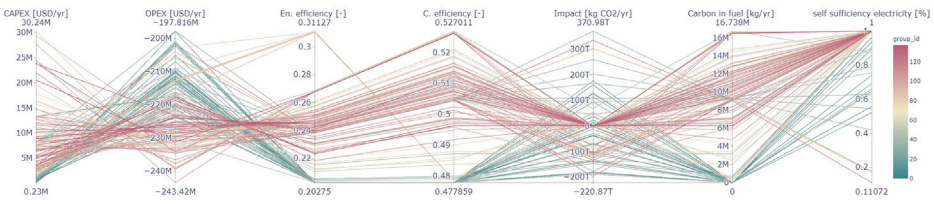


Fig. 8. Unique solutions are displayed in parallel coordinates. Each line represents one solution displayed on multiple axes that can be configuration characteristics, KPIs or objectives. The colour is set by "group_id", which indicates the group of unique solutions.

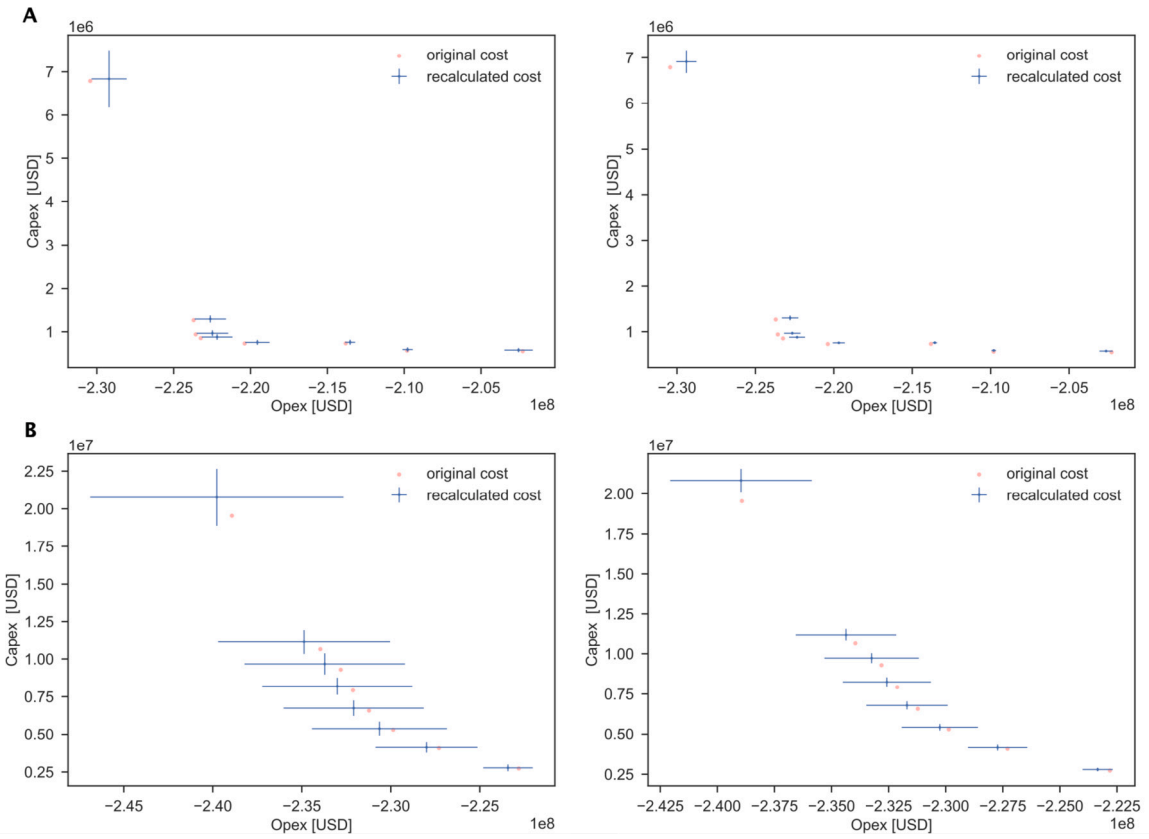


Fig. 9. Original and recalculated cost for 2 samples in N, showing the 95 percentile on the left and the 75 percentile on the right. For sample A, the parameters the optimizer used for generating results are in the mean of the CAPEX, but too optimistic for OPEX. For sample B, the current distributions are more in line with the original parameters used.

Table 2
Economic parameters used for sampling in steering input 1 and 2. Sampling refers to the distribution used in the exploration step (steering input 2). In steering input 1, all parameters are assumed to be quasi-randomly distributed following LHS.

Parameter	Price	Unit	Steering input 1		Steering input 2	Reference
			MIN	MAX	Sampling	
C^{inv}	C_u^{inv}	USD	0.9 $\cdot C_u^{\text{inv}}$	1.1 $\cdot C_u^{\text{inv}}$	Uniform	[76,78]
Interest rate	0.06	%	0.05	0.08	Uniform	
Lifetime	25	years	23	28	Uniform	
Wood price	0.0925	1kg	0.0833	0.1018	Uniform	[80]
Electricity price	0.1016	1 kWh	0.0915	0.1118	Uniform	[81]
Nat gas price	0.0256	USD/kWh	0.0230	0.0282	Uniform	[82]
CO_2 tax	0.0490	USD/kg CO_2eq	0.0441	0.0539	Uniform	[83]
FT fuel price	−1.1080	USD/kg	−1.2188	−0.9972	Uniform	[84]
MeOH price	−0.3900	USD/kg	−0.4290	−0.3510	Uniform	[85]
DME price	−0.8280	USD/kg	−0.9108	−0.7452	Uniform	[86]
Hydrogen price	2.5000	USD/kg	2.2500	2.7500	Uniform	[87]

variable annual operating costs that are dependent on the size and the utilization of the unit $y_{u,t}^{\text{use}}$.

$$\text{OPEX} = \sum_{u}^{\text{U}} \sum_{t=1}^{\text{T}} (c_u^{\text{op1}} \cdot y_{u,t}^{\text{use}} + c_u^{\text{op2}} \cdot f_{u,t}^{\text{mult}}) \cdot t_t^{op} \tag{3}$$

CAPEX: This indicator includes costs associated with the purchase and installation of new equipment C^{inv}, derived from [76,78]. Investment costs are updated to the current year using the CEPCI index of 2019 and annualized over the expected lifetime of the equipment n_u with the interest rate i".

$$\begin{aligned} \text{CAPEX}_u &= \frac{i(1+i)^{n_u}}{(1+i)^{n_u}-1} \cdot C_u^{\text{inv}} \\ \text{CAPEX} &= \sum_u \text{CAPEX}_u \end{aligned} \tag{4}$$

Total Costs: The total annual cost is the sum of OPEX and CAPEX, providing an indication of the plant's profitability.

Payback period: The payback period p_{payback} relates the OPEX and the investment cost made. It needs to be noted that we only include the investment cost related to the biorefinery units in the calculation, and we only consider the fuel selling revenues and resource that can directly be allocated to the fuel production as operating cost.

$$p_{\text{payback}} = \frac{\sum_u C_u^{\text{inv}}}{OPEX^{\text{bio}}} \tag{5}$$

Environmental indicators: In this study, our environmental impact analysis is realized using the Ecoinvent 3.6 database [88] including data on Global Warming Potential (GWP) in [kgCO_2*eq*/year] from IPCC 2013. However, any other indicator can be chosen as well. Further

impact-related KPIs we consider are the amount of carbon and energy stored in the fuel, as well as the total fossil emissions of the mill.

Resilience indicators: We identify KPIs that describe the resilience of a proposed plant configuration. As a measure of resilience, we consider the plant's self sufficiency regarding electricity (E) $\eta_{E,self}$ and fuel consumption $\eta_{F,self}$, measured with the amount of carbon (C) in the fuel generated and bought.

$$\begin{aligned} \eta_{\mathrm{E,self}} &= E_{\mathrm{consume}} - \frac{E_{\mathrm{buy}}}{E_{\mathrm{consume}}} \\ \eta_{F,self} &= C_{\mathrm{fuel,generate}} - \frac{C_{\mathrm{fuel,buy}}}{C_{\mathrm{fuel,generate}}} \end{aligned} \tag{6}$$

System efficiencies: The energetic efficiency η_{en} is an indication of the energetic conversion performance of biomass into liquid fuels. It is calculated using the lower heating values (LHV) of fuels and biomass and the respective mass flows m, as well as the amount of electricity being sold or bought. The carbon conversion efficiency η_{carbon} is defined as the ratio of carbon leaving the system stored in value-added products and the carbon entering with the inputs.

$$\begin{aligned} \eta_{\mathrm{en}} &= \frac{m \cdot LHV_{\mathrm{fuel}} + E_{\mathrm{sell}}}{m \cdot LHV_{\mathrm{biomass,in}} + E_{\mathrm{buy}}} \\ \eta_{\mathrm{carbon}} &= \frac{C_{\mathrm{pulp}} + C_{\mathrm{fuel,generate}}}{C_{\mathrm{pulpwood,in}}} \end{aligned} \tag{7}$$

3.2.2. Decision space exploration, problem formulation and solution generation

As described in the methodology, the economic parameters subjected to uncertainty p_u and the decision variables of the upper level framework Θ are sampled in the steering input 1. $N = 20$ samples are generated in a first step, avoiding taking too much computational time. For each sample $n \in N$, a set of Pareto-optimal solutions is generated. In our case study, two objectives are explored by the lower level framework optimization, OPEX and CAPEX. The relaxed solutions of the worst economic scenario are computed and a tolerance of 20% is accepted, meaning that samples are only used for the calculation of the Pareto-front if the payoff tables are better than 20% of the worst solutions achieved. For each payoff table qualifying for further evaluation, six Pareto points are calculated using Sobol sampling for generation of the ε bounds. According to the criteria defined above, 17 of the 20 payoff tables qualify for further evaluation, leading to a total of 182 function evaluations. The reduced MIP dimensions for a typical problem to be evaluated during the solution generation are 3180 rows, 19374 columns, and 76264 non-zeros. Each function evaluation takes approximately 32 seconds (s), which results in a solution generation time of 97 minutes (min).

3.3. Solution exploration

From the solution generation defined in the previous section, 138 feasible unique solutions are retrieved, satisfying the demand of performing better than the solutions of the worst economic scenario. Following the procedure presented in Fig. 5, all unique solutions are displayed in parallel coordinates (see Fig. 8). For better illustrating the suggested method, we follow the steps of the solution exploration from 2 different DM perspectives and compare the solutions.

3.3.1. Filtering user preferences and analysing uncertainty

Both DMs have no strong preference towards a certain trend in the solutions before analysing the performance under uncertainty; thus all unique solutions are included in the uncertainty calculation. Fig. 9 shows the changes in economic performance for the generated Pareto-front for two samples in N with 75% and 95% percentiles. Depending on the assumed distribution, the solutions perform better or worse economically than in the original optimization result. As a reminder, the economic parameters were also sampled during the generation of solutions for obtaining a diverse solution space. By showing the impact of the economic uncertainty in their solutions to the DMs, they can decide to either take them into account with a given percentile, or to ignore them. In our case study, the impact of the economic uncertainty for the assumed distribution is judged to be quite remarkable by both DMs. Therefore, both include the updated cost in their ranking. DM 1 includes the mean of the updated TOTEX, while DM 2 includes the 95 percentile. Furthermore, the DM 2 discards all solutions with a mean CAPEX above 15 million USD (MUSD), which leaves him with 124 remaining unique solutions.

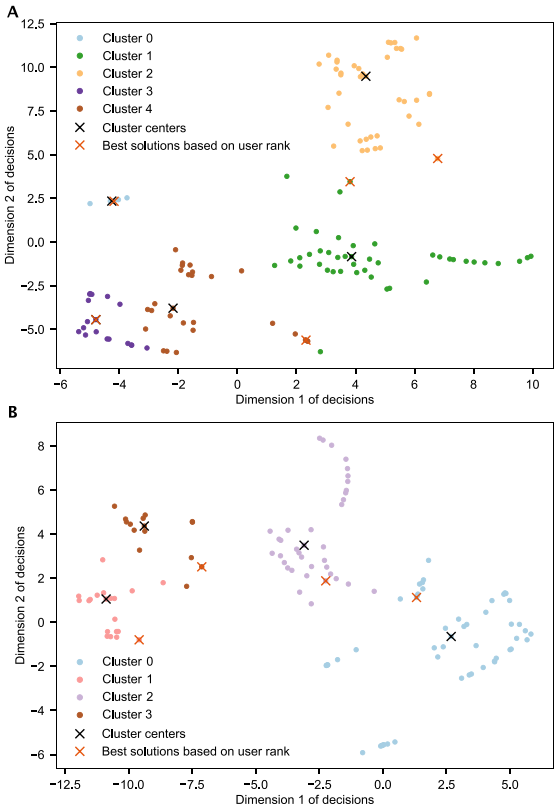


Fig. 10. Cluster results for both DMs, clustering based on optimizer decisions using K-medoids. Best performing solution based on user rank calculations. **A:** Clusters for DM 1, 138 unique solutions included. **B:** Clusters for DM 2, 124 unique solutions included.

Table 3
Decision maker specifications for ranking calculation and resulting user rank ranges.

	TOPSIS specifications			User rank	
	Indices	Weights	Direction	min	max
DM 1	TOTEX mean	1/3	−1		
	$\eta_{\mathrm{E,self}}$	1/3	1		
	η_{Carbon}	1/3	1	0.169	0.845
DM 2	TOTEX 95%	1/2	−1		
	Payback time	1/2	−1	0.335	0.667

3.3.2. Ranking and diversification

For ranking, DM 1 includes mean TOTEX, the carbon efficiency and the self sufficiency of the mill in user rank calculation, all equally weighted. This leads to a user rank between 0.17 and 0.84 for all unique solutions. DM 2 is only interested in TOTEX and payback time, for both the 95 percentile is included, which leads to a user rank between 0.34 and 0.67 (see Table 3). The direction indicates whether a criteria included in TOPSIS should be interpreted as a benefit or as cost.

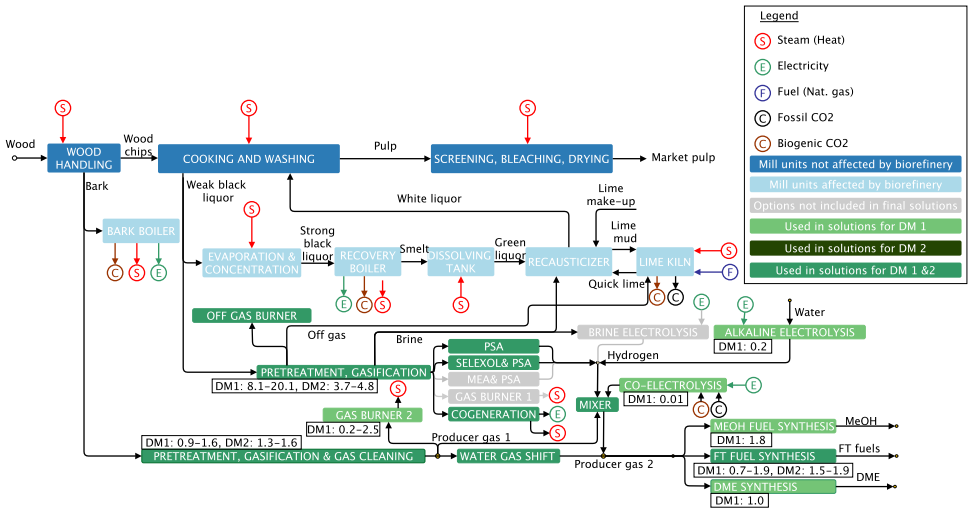


Fig. 11. Result of case study regarding unit choices. f^{mult} for selected units indicated in boxes for both decision makers DM 1 and DM 2.

Clustering the unique solution based on optimizer's decisions leads to five diverse solutions for DM 1, each being the best performing solution based on user rank 1 in the respective cluster. For DM 2 who previously discarded some unique solutions, we get 4 clusters and thus 4 diverse solutions. Fig. 10 displays the generated clusters for both DMs reduced in a two-dimensional space.

For the two solutions sets, the relevant KPIs and user ranks as well as the fuel production are presented in Fig. 12. The associated optimizer's decisions are given in Table 7 in Appendix.

It is noticeable that DM 1 only receives final solutions with fuel production included. Three solutions include FT fuel production, whereas DME and MeOH is picked once each. The three solutions generating FT fuel (cluster 1–3) are differentiated by the method of setting the producer gas ratio to the required one for fuel synthesis (hydrogen production or water gas shift), the origin of the producer gas (dry gasification and co-electrolysis), the valorization of produced hydrogen (used in mixer or in cogeneration) and the gas cleaning mechanisms for hydrogen (Selexol and PSA). The solutions yielding DME and MeOH use both the dry gasification, whereas in the case of DME production, the producer gas ratio is adapted using water gas shift, whereas in the solution yielding MeOH, Alkaline electrolysis additionally used to support the adjustment. The solutions proposed to DM 2 are less diverse regarding the fuel production. Apart from a solution valorizing residuals in the conventional way, FT fuel is produced in three other solutions. When producing FT fuel, the producer gas ratio is adjusted by varying shares of hydrothermal gasification and water gas shift. The solutions proposed to both DMs are visualized in Fig. 11, where the f^{mult} of the major units is included for the selected solutions.

3.3.3. Steering the solution space

The most relevant decisions influencing the calculated user rank are visualized in a polar plot (Fig. 13). Out of the twenty-five possible units to install, three are never chosen and therefore not included in the representation. Particularly influencing the user rank 1 is the size of the water gas shift unit, as well as the recovery boiler, the pretreatment and hydrothermal gasification unit for black liquor. For the solution set 2, the most influencing decisions are the size of the water gas shift unit, the FT synthesis and the gas burner 2. Looking at the bar plots of the decision space exploration in Fig. 13. For both DM 1 and 2, it could be interesting to investigate if the size of the bark treating units, e.g. the dry gasification pretreatment, the bark boiler and the water gas shift unit could be increased, if more bark is made available. All unique solutions as well as the selected solution for DM 1 and 2 reach the bound set to this units which is imposed by the limited amount of bark available in the mill. Adding the option of buying more bark or generating less pulp could generate more interesting results to both DMs. Furthermore, DM 1 should investigate the option of sending more black liquor to hydrothermal gasification, as m_chtg reaches its upper bound for both solution sets. The upper bound in this variable is mainly set due to the feasible unit sizes, however, installing units in parallel could be an option. The recovery boiler is also exploited up to it's maximum capacity for the solutions presented to DM 2. However, this is related to the fact that the recovery boiler is already installed and the maximum capacity displayed is the capacity that is currently installed. Thus, reaching the limit in this case simply means that all black liquor is valorized in the bark boiler. If DM 2 wishes to generate more meaningful solutions, they should investigate the additional purchase of bark, or the option of valorizing more woody biomass in the dry gasification line at the trade-off of less pulp production. Regarding the already generated solution, DM 1 should settle for solution of cluster 1 or 2 regarding the user rank. Looking at the original KPIs determining the user rank, cluster 2 is slightly better and should thus be preferred. According to the user rank, DM 2 should settle for the solution of cluster 0 where no biorefinery units are installed. The next best solution from there is solution of cluster 2, generating FT fuel by gasifying bark.

4. Discussion

Limitations of the methodology concern the exclusive use of independent optimizer decisions for steering and determining the completeness of the solution space, as the correlation between units can bias the displayed information. In this work, this was addressed by only allowing for the sampling and the exploration of unrelated unit sizes. However, knowledge about the unit relations prior to the application of the method is inevitable. For improving this method, sensitivity analysis prior to unit size sampling in the solution generation might be adequate. Another limitation of this work is the absence of uncertainty in the impact factors underlying the analysis. We assumed distribution of economic parameters, but including a similar concept to the environmental assumptions might lead to different results as well. Furthermore, more ranking parameters could be included for guiding the DMs. Nevertheless, our method has proven to enable a comprehensive solution generation exploration for complex optimization problems. The solution space can be explored for a variety of DM preferences

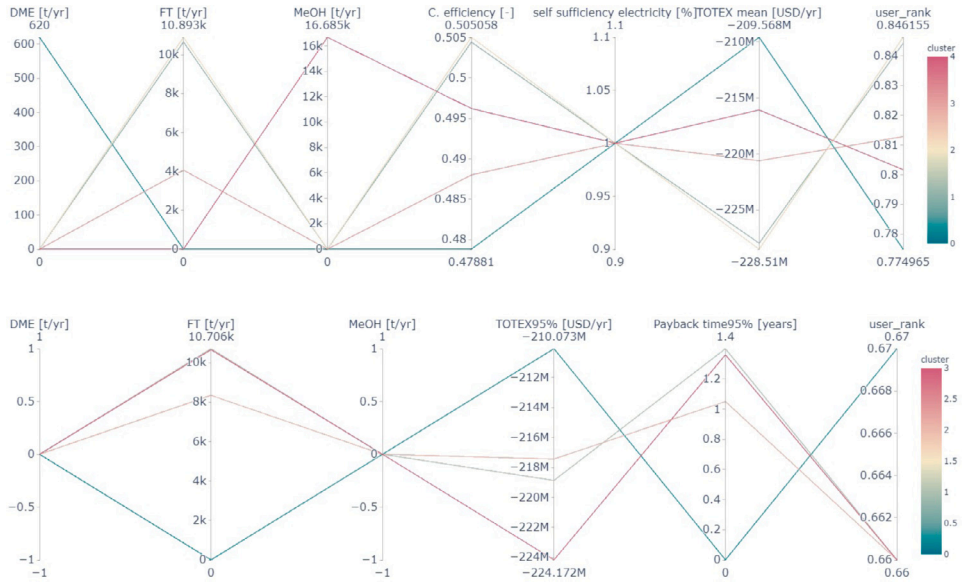


Fig. 12. User rank, involved key performance indicators and annual fuel production displayed for both DMs. Top: DM 1, Bottom: DM 2.

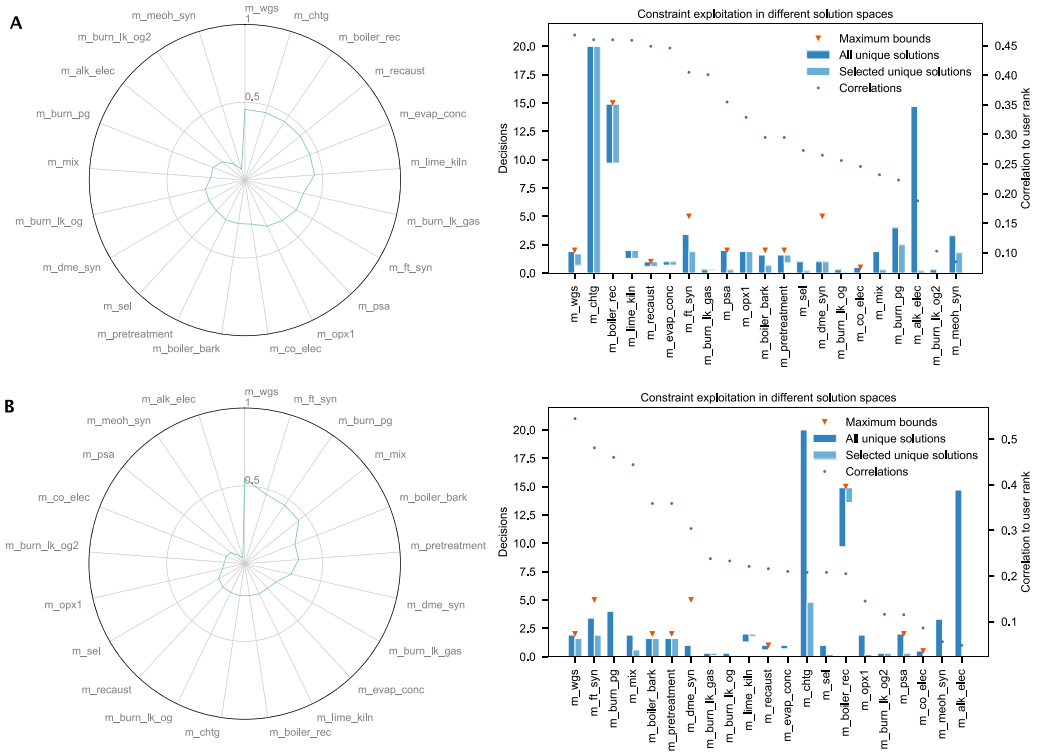


Fig. 13. **A:** DM 1, **B:** DM 2. Spider plots represent the correlation between the optimizer decisions and the user rank, bar plots show the exploitation of the decision space in all unique solutions and the ones selected in the final step.

and specifications, including uncertainty considerations. The step-by-step propagation of the method allows the decision maker to follow the process of solution filtering and to iterate through different steps multiple times to adjust his preferences according to the retrieved results. The practice of visualizing each step gives guidance and makes the process intuitive to understand.

5. Conclusion

The present work addresses the challenge of decision support in complex superstructure optimization by adding a digital twin (InDiT) that consults the decision maker during the solution synthesis and exploration and serves as a translator between human and machine.

The aim hereby is to enable the selection of results not only based on the technical description an engineer provides, but rather to include the decision maker's choices expressed at any time during the process, and thus to provide solutions that respect the decision maker's needs. After defining a superstructure of interest, the steps followed are the definition of steering inputs to capture initial user preferences and the exploration of the decision space, followed by the problem formulation and the solution synthesis. In the solution exploration, a multi-stage filtering is employed, considering user preferences regarding Key Performance Indicators and optimizer decisions, as well as the performance under uncertainty and the diversity of the final solution sets. Furthermore, a steering proposition is provided by the digital twin that suggests how further relevant solutions could be generated. The method was applied for generating solutions for the integration of biorefinery concepts in a typical Kraft pulp mill. Solutions were investigated for two decision maker specifications. In both cases, multiple solutions of comparable performance regarding the decision maker's preferences were retrieved, all of them portraying different system configurations. Furthermore, InDiT was able to suggest how to generate more meaningful solutions. It was shown that the solutions proposed by InDiT are in line with the expressed preferences, and that the suggested method allows for strategic solution generation and holistic exploration of complex optimization problems. We believe that our concept of introducing a digital twin that mirrors the whole decision making process of an engineer and actively communicates with its human counterpart could help the decision making in many complex engineering applications, enabling a holistic understanding of the solution space and the impact of expressed preferences.

CRediT authorship contribution statement

Julia Granacher: Conceptualization, Methodology, Software, Data curation, Formal analysis, Investigation, Validation, Visualization, Writing – original draft, Review & editing. **Tuong-Van Nguyen:** Conceptualization, Methodology, Validation, Writing – review & editing, Supervision. **Rafael Castro-Amoedo:** Conceptualization, Validation, Writing – review & editing. **François Maréchal:** Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgements

The authors would like to thank David Yang Shu for his constructive comments during the synthesis of the presented methodology.

Funding

This research has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 818011 and under the Marie Skłodowska-Curie grant agreements No 754462 and 754354.

Appendix A. Supplementary material on the methodology

Supplementary information on the mathematical formulation of the superstructure model

The methodology for superstructure modeling and optimization is adapted from Gassner and Maréchal [11],Kantor et al. [39]. For each unit u in the system, energy and mass flow models are built to describe the conversion of process streams and physical properties, as well as to obtain the characteristics of the interfaces offered for integration with other units. Depending on the type of unit, flowsheeting models are developed to derive the mass and energy balances of the system. Assuming a set of possible units **U** and a set of possible system states **T**, binary decision variables y_u^{use} and $y_{u,t}^{\text{use}}$ describe, respectively, whether a unit u is installed, and whether it is used in the respective period t. Continuous decision variables f_u^{mult} and $f_{u,t}^{\text{mult}}$ describe the installed size of the unit and the level of usage at which it is operated in each period t, respectively. f_u^{mult} and $f_{u,t}^{\text{mult}}$ are constrained by parametrized upper and lower bounds: $F_u^{\text{min/max}}$. The binaries y_u^{use} and $y_{u,t}^{\text{use}}$ are constrained by the bound Y_u that determines whether a unit is considered for the generation of results. In the superstructure model, those variables are related by the following set of equations (Eq. (8)).

$$\begin{aligned} F_u^{\min} \cdot y_u^{\text{use}} \leq \; & f_u^{\text{mult}} \leq F_u^{\max} \cdot y_u^{\text{use}} \quad \forall\, u \in \mathbf{U} \\ F_u^{\min} \cdot y_{u,t}^{\text{use}} \leq \; & f_{u,t}^{\text{mult}} \leq F_u^{\max} \cdot y_{u,t}^{\text{use}} \quad \forall\, u \in \mathbf{U}, t \in \mathbf{T} \\ & Y_u \geq y_u^{\text{use}} \geq y_{u,t}^{\text{use}} \quad \forall\, u \in \mathbf{U}, t \in \mathbf{T} \end{aligned} \tag{8}$$

Material and energy flow balances are also used to define the unit's requirements. $\dot{m}_{r,u,t}^{+}$ and $\dot{m}_{r,u,t}^{-}$ define the reference mass flowrate of resource r produced and consumed, respectively, in unit u at time step t. Requirements for each resource are satisfied by inside production and imports (Eq. (9)). The overall resource balance ensures that import, export and production are balanced, as formulated in Eq. (10).

The mass balance for each resource layer is closed with the amount of resource r consumed/supplied to the system in time step t ($\dot{M}_{r,j,t}^{-}/\dot{M}_{r,i,t}^{+}$) and balanced by internal needs and connection flowrates of resource r between supplying unit i (**SU**) and consuming unit j (**CU**) given by $\dot{m}_{r,i,j,t}$ (Eqs. (12) and (13)).

$$\sum_{u} f_{u,t}^{\text{mult}} \cdot \dot{m}_{r,u,t}^{+} + \dot{M}_{r,t}^{-} - \sum_{u=1}^{n_u} f_{u}^{\text{mult}} \cdot \dot{m}_{r,u,t}^{-} \geq 0, \quad \forall r \in \mathbf{R}, \forall t \in \mathbf{T} \tag{9}$$

$$\sum_{u=1}^{n_u} f_{u,t}^{\text{mult}} \cdot \dot{m}_{r,u,t}^{+} + \dot{M}_{r,t}^{-} - \dot{M}_{r,t}^{+} - \sum_{u=1}^{n_u} f_{u,t}^{\text{mult}} \cdot \dot{m}_{r,u,t}^{-} = 0, \quad \forall r \in \mathbf{R}, \forall t \in \mathbf{T} \tag{10}$$

$$0 = \sum_{r} f_{u,t}^{\text{mult}} \cdot (\dot{m}_{r,u,t}^{+} - \cdot \dot{m}_{r,u,t}^{-}), \quad \forall u \in \mathbf{U}, \forall t \in \mathbf{T} \tag{11}$$

$$\dot{M}_{r,j,t}^{-} + \sum_{i=1}^{n_i} \dot{m}_{r,i,j,t} = f_{j,t}^{\text{mult}} \cdot \dot{m}_{r,j,t}^{-} \quad \forall r \in \mathbf{R}, \forall j \in \mathbf{CU}, \forall t \in \mathbf{T} \tag{12}$$

$$\dot{M}_{r,i,t}^{+} = f_{i,t}^{\text{mult}} \cdot \dot{m}_{r,i,t}^{+} - \sum_{j=1}^{n_j} \dot{m}_{r,i,j,t} \quad \forall r \in \mathbf{R}, \forall i \in \mathbf{SU}, \forall t \in \mathbf{T} \tag{13}$$

In our approach, all units are connected to a heat exchange system, which allows for the exchange of heat between processes and an external utility system. This ensures that the energy demand is satisfied and that the temperature-enthalpy profile of each unit is considered. Minimum energy requirements are calculated applying the approach presented by Maréchal and Kalitventzeff [111] and based on the work of Linnhoff and Hindmarsh [8]. The energy balance is closed in each temperature interval h (Eq. (14)) and residual heat ($\dot{R}_{t,h}$) flows from higher (h) to lower ($h-1$) temperature (θ) levels. Following thermodynamic feasibility, cascaded heat flows are positive, while values in both the first and the last interval h are zero (Eq. (15)). $\dot{q}_{u,t,h}$ represents the

Table 4
Model parameters and references used in the superstructure. For each unit (bold), main references are given, as well as the main modeling assumptions.

Modeling assumptions	Value	Reference
Pulp mill model		[13,89]
pulp production	1000 adt[1]/day	
black liquor stream	14.89 kg/s	
bark stream available from debarking of wood	67 MW/2000 adt pulp	
Lime kiln and chemical recovery		[90–93]
calcination temp	900° C	
kiln product temperature	300° C	
reburn specific heat	989 J/kg/K	
CO_2 specific heat	919 J/kg/K	
heat of calcination	3270 kJ/kg	
inerts specific heat	1046 J/kg/K	
availability	85%	
Na_2CO_3 in smelt	278.3 g/kg dry solids in black liquor	
Alkaline water electrolysis (AEC)		[94]
water inlet	0.0799 kg/s	
hydrogen production	0.069 kg/s	
oxygen production	0.17 kg/s	
electricity input	1000 kWh	
system efficiency	52 kWe/kg H_2	
Co-electrolyis		[95]
water inlet	1.5335 kg/s	
CO_2 input	2.64 kg/s	
syngas produced	3.01098 kg/s	
oxygen co-produced	1.16254 kg/s	
electricity input	18336 kg/s	
Brine electrolysis		[94,96]
water inlet to dilute brine	0.93 kg/s	
$NaCo_3$ input	0.093 kg/s	
$NaOH$ output	0.019 kg/s	
$NaCo_3$ output	0.065 kg/s	
electricity input	1000 kg/s	
hydrogen production	0.0005 kg/s	
Pretreatment and gasification of black liquor		
Hydrolysis		[63,70,97]
lignin fraction in organic biomass	94 $wt\%_{daf}$[a]	
H/C-ratio of lignin	1.11 mol_{daf}	
O/C-ratio of lignin	0.33 mol_{daf}	
effective water content	93 wt%	
Decomposition of carboxylic salts		[63,70,97,98]
reactor yield	70%	
Salt separator		[63,70,97,99–101]
recovery of inorganic cooking chemicals	100%	
organic loss in salt brine	10%	
Gasification and expansion		[63,97,102]
reactor temperature	700° C	
reactor pressure	250 bar	
gas expander isentropic efficiency	80%	
liquid expander isentropic efficiency	82%	
Pressure Swing Adsorption (PSA)		[103,104]
recovery rate H_2	52%	
H_2 purity	99.996 mol%	
number of beds	4	
Selexol and PSA		[63,68,105,106]
recovery rate CO_2 Selexol	4%	
recovery rate H_2 Selexol	100%	
recovery rate H_2 PSA	80%	
number of beds PSA	6	
H_2 purity	99.99+ mol%	
Monoethanolamine (MEA) and PSA		[68,107]
recovery CO_2 MEA	5	
recovery H_2 PSA	80	
number of beds PSA	6	
H_2 purity	99.99+ mol%	
Pretreatment of bark, gasification, gas cleaning,WGS, fuel synthesis		[12,107–110]
drying technology	steam drying	

(*continued on next page*)

Table 4 (*continued*).

Modeling assumptions	Value	Reference
Pulp mill model		[13,89]
moisture content after drying	10%	
gasification technology	Directly heated entrained flow (EF)	
operating conditions of gasification	1350 °C, 30 bar	
cold gas cleaning temperature	150 °C	
WGS temperature	300 °C	
CO_2 removal	MEA	
FT synthesis	25 bar, 220 °C	
DME synthesis	50 bar, 277 °C	
MeOH synthesis	85 bar, 315 °C	
upgrading FT/DME/MeOH	private data/ distillation /distillation	

[a]daf: dry ash-free

reference heat load for unit u in time step t and temperature interval h.

$\forall h \in \mathbf{H}$ with $\theta_{h+1} \geq \theta_h$

$$\sum_{u=1}^{n_u} \dot{q}_{u,t,h} \cdot f_{u,t}^{\text{mult}} + \dot{R}_{t,h+1} - \dot{R}_{t,h} = 0 \quad \forall t \in \mathbf{T} \tag{14}$$

$$\dot{R}_{t,h} \geq 0, \quad \dot{R}_{t,1} = \dot{R}_{t,nh+1} = 0 \quad \forall t \in \mathbf{T} \tag{15}$$

Chemical reactions are included in the respective units by calling simulation software to provide reaction characteristics and the corresponding energy and mass flows. For the economic analysis, the operating and investment costs are calculated as a function of equipment size, using cost functions available in literature. For the environmental assessment, the LCI (Life Cycle Inventory) Ecoinvent database [88] is used to estimate the environmental impacts associated with waste streams and material use for the unit construction [112].

Supplementary information on the implementation of the solution exploration

Our decision exploration is performed using *streamlit*, an open-source app framework for Machine Learning and Data Science in Python [113] in combination with the *pandas* library for data analysis [114] and other python libraries and packages. The main implementation steps are specified further in the following paragraph. For displaying the solutions in parallel coordinates, the plotly express library is used [115]. When calculating the distribution of key performance indicators for the uncertainty distributions, the equations in are used with the respective distribution of parameters as given in Table 2. For getting the desired percentile of the distribution and other descriptive statistics, we use the *pandas describe* function [114]. Calculating the rank of the solution with TOPSIS, we use the the *topsis-jamesfallon* module [116], which uses a collection of criteria with associated weights and directions as given in Table 3 to calculate the user rank. More information about the mathematical specification of the TOPSIS calculation can be found in [117]. For steering the solution generation from the generated results, the correlation between each KPI of interest and the optimizer decision is obtained using the *pandas corr* function [114] with the spearman rank correlation coefficient [118].

Supplementary material on the case study

Supplementary material on the process superstructure

Table 4 gives an overview of the main modeling assumptions included in the process superstructure and the references the reader may consult for further information for the process model details.

For the industrial Kraft mill studied in this work, the typical pulp production uses 60% of hardwood species (poplar, beech, oak and chestnut) and 40% of softwood species (pines and spruce) by weight. Table 5 shows the biomass composition (on a dry, ash-free basis) used in the analysis.

Table 5
Composition of bark biomass on a dry ash-free basis.

Feedstock	Softwood bark	Hardwood bark	Typical 60/40 Mix
C %	50.9%	51.1%	51.1%
H %	6.1%	6.3%	6.2%
N %	0.4%	0.4%	0.4%
S %	0.1%	0.1%	0.1%
O %	42.5%	42.2%	38.0%

Table 6
Composition of Softwood Black Liquor Solids on a dry ash-free (*daf*) basis.

Parameter	Value	Unit
Lignin	31	$wt\%_{daf}$
Poly-saccharides	2	$wt\%_{daf}$
Aliphatic salts[a]	40	$wt\%_{daf}$
Inorganics in black liquor solids	22	$wt\%_{daf}$

[a]Includes aliphatic carboxylic acids ($29wt\%_{daf}$) and sodium bound to organics ($11wt\%_{daf}$).

The weak black liquor is assumed to contain 18wt% dry solids (corresponding to a water content of 82 wt%). In this study, as per [70] the composition of black liquor is restricted to the following species: lignin, poly-saccharides, aliphatic (carboxylic) salts and inorganics. The inorganics are modelled by four main compounds, namely sodium carbonate Na_2CO_3, sodium sulphate Na_2SO_4, sodium hydroxide NaOH and sodium sulphide Na_2S, with relative proportions as per [70]. Extractives are not included in this analysis as they are typically removed in the weak black liquor storage tanks. It is assumed that any extractives remaining in the BL stream exit the system as losses in the salt separator and therefore do not affect the syngas yield. The assumed composition is displayed in Table 6.

Supplementary information on the results

Table 7 shows the decisions of the selected solutions for both DMs analyzed in the case study.

Table 7
Decisions of selected solutions for both DMs.

	DM 1					DM 2			
Decision f^{mult} / Solution	0	1	2	3	4	0	1	2	3
m_pretreatment	0.9	1.6	1.6	1.6	1.6	0	1.6	1.3	1.6
m_wgs	1	1.6	1.6	0.7	1.7	0	1.3	1.5	1.6
m_meoh_syn	0	0	0	0	1.8	0	0	0	0
m_dme_syn	1	0	0	0	0	0	0	0	0
m_ft_syn	0	1.9	1.9	0.7	0	0	1.9	1.5	1.9
m_chtg	0	8.1	20	0	0	0	4.8	0	3.7
m_psa	0	0	0.3	0	0	0	0.3	0	0
m_mea_psa	0	0	0	0	0	0	0	0	0
m_sel	0	0.2	0	0	0	0	0.2	0	0.2
m_mix	0	0.3	0.3	0	0	0	0.6	0	0.3
m_br_elec	0	0	0	0	0	0	0	0	0
m_alk_elec	0	0	0	0	0.2	0	0	0	0
m_co_elec	0	0	0.01	0	0	0	0	0	0
m_opx1	0	0.7	1.9	0	0	0	0	0	0.2
m_burn_1	0	0	0	0	0	0	0	0	0
m_burn_pg	0	0	0	2.5	0.2	0	0	0	0
m_lime_kiln	2	1.7	1.3	2	2	2	1.8	2	1.9
m_burn_lk_og	0	0.1	0	0	0	0	0.4	0	0.1
m_burn_lk_gas	0.3	0.2	0.2	0.3	0.3	0.3	0.1	0.3	0.2
m_recaust	1	0.8	0.6	1	1	1	0.9	1	0.9
m_evap_conc	1	0.9	0.7	1	1	1	0.9	1	0.9
m_boiler_rec	14.9	12.8	9.7	14.9	14.9	14.9	13.6	14.9	13.9
m_boiler_bark	0.7	0	0	0	0	1.6	0	0.3	0

References

[1] Mencarelli L, Chen Q, Pagot A, Grossmann IE. A review on superstructure optimization approaches in process system engineering. Comput Chem Eng 2020;136:106808. http://dx.doi.org/10.1016/j.compchemeng.2020.106808.

[2] Stephanopoulos G, Reklaitis GV. Process systems engineering: From solvay to modern bio- and nanotechnology.: A history of development, successes and prospects for the future. In: Multiscale simulation. Chem Eng Sci 2011;66(19):4272–306. http://dx.doi.org/10.1016/j.ces.2011.05.049.

[3] Henao CA, Maravelias CT. Surrogate-based superstructure optimization framework. AIChE J 2011;57(5):1216–32. http://dx.doi.org/10.1002/aic.12341.

[4] Cozad A, Sahinidis NV, Miller DC. Learning surrogate models for simulation-based optimization. AIChE J 2014;60(6):2211–27. http://dx.doi.org/10.1002/aic.14418.

[5] Rudd DF, Powers GJ, Siirola JJ. Process synthesis. Prentice-Hall; 1973.

[6] Douglas JM. Conceptual design of chemical processes. McGraw-Hill: M6JTAAAAMAAJ; 1988, Google-Books-ID.

[7] Ciric AR, Mumtaz HS, Corbett G, Reagan M, Seider WD, Fabiano LA, Kolesar DM, Widagdo S. Azeotropic distillation with an internal decanter. Comput Chem Eng 2000;24(11):2435–46. http://dx.doi.org/10.1016/S0098-1354(00)00603-7.

[8] Linnhoff B, Hindmarsh E. The pinch design method for heat exchanger networks. Chem Eng Sci 1983;38(5):745–63, 00000. http://dx.doi.org/10.1016/0009-2509(83)80185-7.

[9] Grossmann IE. Mixed-integer programming approach for the synthesis of integrated process flowsheets. In: Papers from the 25th conicet international conference. Comput Chem Eng 1985;9(5):463–82. http://dx.doi.org/10.1016/0098-1354(85)80023-5.

[10] Liu P, Nguyen T-D, Cai X, Jiang X. Finding multiple optimal solutions to optimal load distribution problem in hydropower plant. Energies 2012;5(5):1413–32. http://dx.doi.org/10.3390/en5051413.

[11] Gassner M, Maréchal F. Methodology for the optimal thermo-economic, multi-objective design of thermochemical fuel production from biomass. In: Selected papers from the 17th european symposium on computer aided process engineering held in bucharest, romania, may 2007. Comput Chem Eng 2009;33(3):769–81. http://dx.doi.org/10.1016/j.compchemeng.2008.09.017.

[12] Celebi AD, Sharma S, Ensinas AV, Maréchal F. Next generation cogeneration system for industry – combined heat and fuel plant using biomass resources. Chem Eng Sci 2019;204:59–75. http://dx.doi.org/10.1016/j.ces.2019.04.018.

[13] Kermani M, Kantor ID, Wallerand AS, Granacher J, Ensinas AV, Maréchal F. A holistic methodology for optimizing industrial resource efficiency. Energies 2019;12(7):1315. http://dx.doi.org/10.3390/en12071315.

[14] Amoedo RC, Damartzis T, Granacher J, Marechal FMA. System design and performance evaluation of wastewater treatment plants coupled with hydrothermal liquefaction and gasification. Front Energy Res 2020;8. http://dx.doi.org/10.3389/fenrg.2020.568465.

[15] Fazlollahi S, Schüler N, Maréchal F. A solid thermal storage model for the optimization of buildings operation strategy. Energy 2015;88:209. http://dx.doi.org/10.1016/j.energy.2015.04.085.

[16] Suciu R-A, Girardin L, Maréchal F. Energy integration of CO2 networks and Power to Gas for emerging energy autonomous cities in Europe. Energy 2018;157:830. http://dx.doi.org/10.1016/j.energy.2018.05.083.

[17] Meignan D, Knust S, Frayret J-M, Pesant G, Gaud N. A review and taxonomy of interactive optimization methods in operations research. Trans Interact Intell Syst 2015;5:1–43. http://dx.doi.org/10.1145/2808234.

[18] Cajot S, Schüler N, Peter M, Koch A, Maréchal F. Interactive optimization with parallel coordinates: Exploring multidimensional spaces for decision support. Front ICT 2019;5. http://dx.doi.org/10.3389/fict.2018.00032.

[19] Branke J, Branke J, Deb K, Miettinen K, Slowiński R. Multiobjective optimization: Interactive and evolutionary approaches. Springer Science & Business Media; 2008.

[20] Balas E, Jeroslow R. Canonical cuts on the unit hypercube. SIAM J Appl Math 1972;23(1):61–9. http://dx.doi.org/10.1137/0123007.

[21] Sahinidis NV, Grossmann IE, Fornari RE, Chathrathi M. Optimization model for long range planning in the chemical industry. Comput Chem Eng 1989;13(9):1049–63. http://dx.doi.org/10.1016/0098-1354(89)87046-2.

[22] Trutnevyte E. Does cost optimization approximate the real-world energy transition? Energy 2016;106:182–93. http://dx.doi.org/10.1016/j.energy.2016.03.038.

[23] DeCarolis JF. Using modeling to generate alternatives (MGA) to expand our thinking on energy futures. Energy Econ 2011;33(2):145–52. http://dx.doi.org/10.1016/j.eneco.2010.05.002.

[24] Hennen M, Lampe M, Voll P, Bardow A. SPREAD – Exploring the decision space in energy systems synthesis. In: ESCAPE-26, Comput Chem Eng In: ESCAPE-26, 2017;106:297–308.http://dx.doi.org/10.1016/j.compchemeng.2017.06.002,

[25] Liu P, Cai X, Guo S. Deriving multiple near-optimal solutions to deterministic reservoir operation problems. Water Resour Res - WATER RESOUR RES 2011;47. http://dx.doi.org/10.1029/2011WR010998.

[26] Cajot S, Mirakyan A, Koch A, Maréchal F. Multicriteria decisions in urban energy system planning: A review. Front Energy Res 2017;5. http://dx.doi.org/10.3389/fenrg.2017.00010.

[27] Liu J, Dwyer T, Marriott K, Millar J, Haworth A. Understanding the relationship between interactive optimisation and visual analytics in the context of prostate brachytherapy. IEEE Trans Vis Comput Graphics 2018;24(1):319–29. http://dx.doi.org/10.1109/TVCG.2017.2744418.

[28] Shenfield A, Fleming PJ, Alkarouri M. Computational steering of a multi-objective evolutionary algorithm for engineering design. Eng Appl Artif Intell 2007;20(8):1047–57. http://dx.doi.org/10.1016/j.engappai.2007.01.005.

[29] Abi Akle A, Minel S, Yannou B. Information visualization for selection in Design by Shopping. Res Eng Des 2017;28(1):99–117. http://dx.doi.org/10.1007/s00163-016-0235-2.

[30] Miettinen K. Survey of methods to visualize alternatives in multiple criteria decision making problems. OR Spectrum 2014.

[31] Moret S. Strategic energy planning under uncertainty. Infoscience 2017. Number: THESIS Publisher: EPFL. http://dx.doi.org/10.5075/epfl-thesis-7961.

[32] Soroudi A, Amraee T. Decision making under uncertainty in energy systems: State of the art. Renew Sustain Energy Rev 2013;28:376–84. http://dx.doi.org/10.1016/j.rser.2013.08.039.

[33] Zhou P, Ang BW, Poh KL. Decision analysis in energy and environmental modeling: An update. Energy 2006;31(14):2604–22.

[34] Babonneau F, Vial J-P, Apparigliato R. Robust optimization for environmental and energy planning. In: Filar JA, Haurie A, editors. Uncertainty and environmental decision making: A handbook of research and best practice. International series in operations research & management science, Boston, MA: Springer US; 2010, p. 79–126. http://dx.doi.org/10.1007/978-1-4419-1129-2_3.

[35] Ben-Tal A, Nemirovski A. Robust solutions of uncertain linear programs. Oper Res Lett 1999;25(1):1–13. http://dx.doi.org/10.1016/S0167-6377(99)00016-4.

[36] Tao F, Zhang H, Liu A, Nee AY. Digital twin in industry: State-of-the-art. IEEE Trans Ind Inf 2019;15(4):2405–15. http://dx.doi.org/10.1109/TII.2018.2873186.

[37] IBM. Cheat sheet: What is digital twin? Internet of Things blog. IBM Bus Oper Blog 2020. URL https://www.ibm.com/blogs/internet-of-things/iot-cheat-sheet-digital-twin/.

[38] Semeraro C, Lezoche M, Panetto H, Dassisti M. Digital twin paradigm: A systematic literature review. Comput Ind 2021;130:103469. http://dx.doi.org/10.1016/j.compind.2021.103469.

[39] Kantor I, Robineau J-L, Bütün H, Maréchal F. A mixed-integer linear programming formulation for optimizing multi-scale material and energy integration. Front Energy Res 2020;8:49. http://dx.doi.org/10.3389/fenrg.2020.00049.

[40] Fourer R, Gay D, Kernighan B. AMPL: A Modeling language for mathematical programming. 2nd ed. Duxbury-Thomson: Management Science - MANAGE SCI Publication Title: Management Science - MANAGE SCI; 2002, http://dx.doi.org/10.1287/mnsc.36.5.519, Journal Abbreviation.

[41] IBM. CPLEX User's Manual. Technical report, IBM ILOG CPLEX Optimization Studio; 2017, p. 586, URL https://www.ibm.com/docs/en/SSSA5P_12.7.1/ilog.odms.studio.help/pdf/usrcplex.pdf.

[42] Kermani M. Methodologies for simultaneous optimization of heat, mass, and power in industrial processes (Ph.D. thesis), Switzerland: EPFL; 2018, URL https://infoscience.epfl.ch/record/257269?p=Kermani.

[43] McKay MD, Beckman RJ, Conover WJ. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 1979;21(2):239–45. http://dx.doi.org/10.2307/1268522.

[44] Haimes Y, Lasdon L, Wismer D. On a bicriterion formulation of the problems of integrated system identification and system optimization. IEEE Trans Syst Man Cybern 1971;SMC-1(3):296–7. http://dx.doi.org/10.1109/TSMC.1971.4308298, Conference Name: IEEE Transactions on Systems, Man, and Cybernetics.

[45] Mavrotas G. Effective implementation of the ϵ-constraint method in multi-objective mathematical programming problems. Appl Math Comput 2009;213(2):455–65. http://dx.doi.org/10.1016/j.amc.2009.03.037.

[46] Wierzbicki AP. The need for and possible methods of objective ranking. In: Ehrgott M, Figueira JR, Greco S, editors. Trends in multiple criteria decision analysis. International series in operations research & management science, Boston, MA: Springer US; 2010, p. 37–56. http://dx.doi.org/10.1007/978-1-4419-5904-1_2.

[47] Burhenne S, Jacob D, Henze GP. Sampling based on Sobol sequences for Monte Carlo techniques applied to building simulations. In: Proceedings of building simulation 2011, 8. Sydney; 2011.

[48] Sobol IM. The distribution of points in a cube and the approximate evaluation of integrals. Zh Vychisl Mat I Mat Fiz 1967;7(4):784–802.

[49] Copado-Méndez PJ, Pozo C, Guillén Gosalbez G, Jiménez L. Enhancing the ϵ-constraint method through the use of objective reduction and random sequences: Application to environmental problems. Comput Chem Eng 2016;87:36–48, Accepted: 2020-08-24T09:20:37Z. http://dx.doi.org/10.1016/j.compchemeng.2015.12.016.

[50] Celebi AD, Ensinas AV, Sharma S, Maréchal F. Early-stage decision making approach for the selection of optimally integrated biorefinery processes. Energy 2017;137:908–16. http://dx.doi.org/10.1016/j.energy.2017.03.080.

[51] Inselberg A. Multidimensional detective. In: Proceedings of VIZ '97: visualization conference, information visualization symposium and parallel rendering symposium. Phoenix, AZ, USA: IEEE Comput. Soc; 1997, p. 100–7. http://dx.doi.org/10.1109/INFVIS.1997.636793.

[52] Li M, Zhen L, Yao X. How to read many-objective solution sets in parallel coordinates [educational forum]. IEEE Computat Intell Mag 2017;12(4):88–100. http://dx.doi.org/10.1109/MCI.2017.2742869.

[53] Bandaru S, Ng AHC, Deb K. Data mining methods for knowledge discovery in multi-objective optimization: Part A - Survey. Expert Syst Appl 2017;70:139–59. http://dx.doi.org/10.1016/j.eswa.2016.10.015.

[54] Heinrich J, Weiskopf D. State of the art of parallel coordinates. Eurographics 2013 - state of the art reports, 2012, p. 22. http://dx.doi.org/10.2312/CONF/EG2013/STARS/095-116.

[55] Tock L, Maréchal F. Decision support for ranking Pareto optimal process designs under uncertain market conditions. Comput Chem Eng 2015;83:165. http://dx.doi.org/10.1016/j.compchemeng.2015.06.009.

[56] Hwang C-L, Yoon K. Multiple attribute decision making: methods and applications a state-of-the-art survey. Lecture notes in economics and mathematical systems, Berlin Heidelberg: Springer-Verlag; 1981, http://dx.doi.org/10.1007/978-3-642-48318-9.

[57] Papathanasiou J, Ploskas N. TOPSIS. In: Papathanasiou J, Ploskas N, editors. Multiple criteria decision aid : Methods, examples and python implementations. Springer optimization and its applications, Cham: Springer International Publishing; 2018, p. 1–30. http://dx.doi.org/10.1007/978-3-319-91648-4_1.

[58] Aguirre O, Taboada H. A clustering method based on dynamic self organizing trees for post-Pareto optimality analysis. In: Complex adaptive sysytems. Procedia Comput Sci 2011;6:195–200. http://dx.doi.org/10.1016/j.procs.2011.08.037.

[59] Kaufman L, Rouseeuw PJ. Finding groups in data: an introduction to cluster analysis. Wiley; 1990, https://www.wiley.com/en-us/Finding+Groups+in+Data%3A+An+Introduction+to+Cluster+Analysis-p-9780471735786.

[60] Park H-S, Jun C-H. A simple and fast algorithm for K-medoids clustering. Expert Syst Appl 2009;36(2, Part 2):3336–41. http://dx.doi.org/10.1016/j.eswa.2008.01.039.

[61] Satopaa V, Albrecht J, Irwin D, Raghavan B. Finding a "kneedle" in a haystack: Detecting knee points in system behavior. In: 2011 31st international conference on distributed computing systems workshops. Minneapolis, MN, USA: IEEE; 2011, p. 166–71. http://dx.doi.org/10.1109/ICDCSW.2011.20.

[62] Viana Ensinas A, Codina Gironès V, Queiroz Albarelli J, Maréchal F, Silva MA. Thermo-economic optimization of integrated first and second generation sugarcane ethanol plant. Chem Eng Trans 2013. http://dx.doi.org/10.3303/CET1335087.

[63] Gassner M, Vogel F, Heyen G, Maréchal F. Optimal process design for the polygeneration of SNG, power and heat by hydrothermal gasification of waste biomass: Thermo-economic process modelling and integration. Energy Environ Sci 2011;4(5):1726–41. http://dx.doi.org/10.1039/C0EE00629G.

[64] Santibañez Aguilar JE, González-Campos JB, Ponce-Ortega JMa, Serna-González M, El-Halwagi MM. Optimal planning of a biomass conversion system considering economic and environmental aspects. Ind Eng Chem Res 2011;50(14):8558–70. http://dx.doi.org/10.1021/ie102195g.

[65] Ng RTL, Tay DHS, Ng DKS. Simultaneous process synthesis, heat and power integration in a sustainable integrated biorefinery. Energy Fuels 2012;26(12):7316–30. http://dx.doi.org/10.1021/ef301283c.

[66] Mongkhonsiri G, Charoensuppanimit P, Anantpinijwatna A, Gani R, Assabumrungrat S. Process development of sustainable biorefinery system integrated into the existing pulping process. J Cleaner Prod 2020;255:120278. http://dx.doi.org/10.1016/j.jclepro.2020.120278.

[67] Naqvi M, Yan J, Fröling M. Bio-refinery system of DME or CH4 production from black liquor gasification in pulp mills. Bioresour Technol 2009;101(3):937–44. http://dx.doi.org/10.1016/j.biortech.2009.08.086.

[68] Cao C, Guo L, Jin H, Cao W, Jia Y, Yao X. System analysis of pulping process coupled with supercritical water gasification of black liquor for combined hydrogen, heat and power production. Energy 2017;132. http://dx.doi.org/10.1016/j.energy.2017.05.104.

[69] Granacher J, Maréchal F, Celebi AD, Kermani M. Potential of hydrothermal black liquor gasification integrated in pulp production plant. In: Proceedings of ECOS 2019. 2019, p. 2299–309.

[70] Magdeldin M, Järvinen M. Supercritical water gasification of Kraft black liquor: Process design, analysis, pulp mill integration and economic evaluation. Appl Energy 2020;262:114558. http://dx.doi.org/10.1016/j.apenergy.2020.114558.

[71] Consonni S, Katofsky RE, Larson ED. A gasification-based biorefinery for the pulp and paper industry. In: Special issue on biorefinery integration. Chem Eng Res Des 2009;87(9):1293–317. http://dx.doi.org/10.1016/j.cherd.2009.07.017.

[72] Andersson J, Lundgren J, Marklund M. Methanol production via pressurized entrained flow biomass gasification – techno-economic comparison of integrated vs. stand-alone production. Biomass Bioenergy 2014;64:256–68. http://dx.doi.org/10.1016/j.biombioe.2014.03.063.

[73] International Energy Agency. European Union 2020-energy policy review. Technical Report, International Energy Agency; 2020, URL https://iea.blob.core.windows.net/assets/ec7cc7e5-f638-431b-ab6e-86f62aa5752b/European_Union_2020_Energy_Policy_Review.pdf.

[74] Naqvi M, Yan J, Dahlquist E. Black liquor gasification integrated in pulp and paper mills: A critical review. Bioresour Technol 2010;101(21):8001–15. http://dx.doi.org/10.1016/j.biortech.2010.05.013.

[75] Belsim. VALI | belsim - process data validation and reconciliation experts. 2021, URL http://www.belsim.com/vali.

[76] Turton R, Bailie RC, Whiting WB, Shaeiwitz JA. Analysis, synthesis and design of chemical processes. Pearson Education: kWXyhVXztZ8C; 2008, Google-Books-ID.

[77] AACE International. 18R-97: Cost Estimate Classification System - As Applied in Engineering, Procurement, and Construction for the Process Industries. Technical report, AACE, Inc.; 2005, p. 10.

[78] Ulrich GD, Vasudevan PT. Chemical engineering process design and economics: A practical guide. 2nd ed. CRC Press; 2003.

[79] Celebi AD. Methodology for the identification of promising integrated biorefineries. Infoscience 2019. http://dx.doi.org/10.5075/epfl-thesis-9666, URL https://infoscience.epfl.ch/record/271315.

[80] United Nations, Food and Agriculture Organization of the United Nations. Forest products annual market review, 2018–2019. Technical report ECE/TIM/SP/48, United Nations; 2019, Issue: ECE/TIM/SP/48. URL https://www.unece.org/forests/areas-of-work/forestsfpm/outputs/forestsfpamr/forestsfpamr2019.html.

[81] Eurostat. Electricity price statistics. Eurostat 2021. URL https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Electricity_price_statistics.

[82] European Commission. Energy prices and costs in Europe. Technical report, European Commission; 2019, URL https://ec.europa.eu/energy/sites/ener/files/epc_report_final_1.pdf.
[83] Asen E. Carbon taxes in europe. Tax Foundation 2020. Section: Global Tax Maps. URL https://taxfoundation.org/carbon-taxes-in-europe-2020/.
[84] Landälv I, Maniatis K, Heuvel Evd, Kalligeros S, Waldheim L. Building up the future, cost of biofuel : sub group on advanced biofuels : sustainable transport forum. 2018, URL http://op.europa.eu/en/publication-detail/-/publication/13e27082-67a2-11e8-ab9c-01aa75ed71a1.
[85] Methanex. Methanex monthly average regional posted contract price history. Technical report, Methanex; 2021, URL https://www.methanex.com/our-business/pricing.
[86] Zeman P, Hönig V, Procházka P, Mařík J. Dimethyl ether as a renewable fuel for diesel engines. Agron Res 2017;15. http://dx.doi.org/10.15159/ar.17.067, Accepted: 2017-10-25T07:07:18Z.
[87] Ball M, Weeda M. 11 - The hydrogen economy—Vision or reality? In: Ball M, Basile A, Veziroğlu TN, editors. Compendium of hydrogen energy. Woodhead publishing series in energy, Oxford: Woodhead Publishing; 2016, p. 237–66. http://dx.doi.org/10.1016/B978-1-78242-364-5.00011-7.
[88] Wernet G, Bauer C, Steubing B, Reinhard J, Moreno-Ruiz E, Weidema B. The ecoinvent database version 3 (part I): overview and methodology. Int J Life Cycle Assess 2016;21(9):1218–30. http://dx.doi.org/10.1007/s11367-016-1087-8.
[89] Pettersson K, Harvey S. Comparison of black liquor gasification with other pulping biorefinery concepts — Systems analysis of economic performance and CO2 emissions. In: 7th Biennial international workshop “advances in energy studies”. Energy 2012;37(1):136–53. http://dx.doi.org/10.1016/j.energy.2011.10.020.
[90] Adams TN. Lime kiln principles and operations. Technical report, 2007, p. 15.
[91] Sanchez DR. Recausticizing - Principles and practice. In: Kraft recovery short course. Orlando, FL: TAPPI Press; 2000, p. 30.
[92] Castro AdS, Figueiredo LS. Optimization of lime kilns based on strategies of advanced process control - case study Cenibra. In: 5 Th International colloquium on eucalyptus pulp. Bahia, Brazil. 2011. p. 10.
[93] Gullichsen J, Paulapuro H, Yhdistys SP-I, Technical Association of the Pulp and Paper Industry, editors. Papermaking science and technology: a series of 19 books covering the latest technology and future trends. Book 6 B: Chemical pulping [...]. Helsinki: Fapet Oy; 1999, OCLC: 833680723.
[94] Matute G, Yusta J, Correas L. Techno-economic modelling of water electrolysers in the range of several MW to provide grid services while generating hydrogen for different applications: A case study in Spain applied to mobility with FCEVs. Int J Hydrogen Energy 2019;44(33):17431–42. http://dx.doi.org/10.1016/j.ijhydene.2019.05.092.
[95] Zhang X, Song Y, Wang G, Bao X. Co-electrolysis of CO2 and H2O in high-temperature solid oxide electrolysis cells: Recent advance in cathodes. In: CO2 capture storage and utilization. J Energy Chemistry 2017;26(5):839–53. http://dx.doi.org/10.1016/j.jechem.2017.07.003.
[96] Simon A, Fujioka T, Price WE, Nghiem LD. Sodium hydroxide production from sodium carbonate and bicarbonate solutions using membrane electrolysis: A feasibility study. Separ Purif Technol 2014;127:70–6. http://dx.doi.org/10.1016/j.seppur.2014.02.020.
[97] Mian A. Optimal design methods applied to solar-assisted hydrothermal gasification plants (Ph.D. thesis), Switzerland: EPFL; 2016, URL https://infoscience.epfl.ch/record/216954?p=Alberto+Mian.
[98] Onwudili JA, Williams PT. Hydrothermal reactions of sodium formate and sodium acetate as model intermediate products of the sodium hydroxide-promoted hydrothermal gasification of biomass. Green Chem 2010;12(12):2214. http://dx.doi.org/10.1039/c0gc00547a.
[99] Schubert M, Regler JW, Vogel F. Continuous salt precipitation and separation from supercritical water. Part 1: Type 1 salts. J Supercrit Fluids 2010;52(1):99–112. http://dx.doi.org/10.1016/j.supflu.2009.10.002.
[100] Schubert M, Aubert J, Müller JB, Vogel F. Continuous salt precipitation and separation from supercritical water. Part 3: Interesting effects in processing type 2 salt mixtures. J Supercrit Fluids 2012;61:44–54. http://dx.doi.org/10.1016/j.supflu.2011.08.011.
[101] Luterbacher JS, Fröling M, Vogel F, Maréchal F, Tester JW. Hydrothermal gasification of waste biomass: Process design and life cycle asessment. Environ Sci Technol 2009;43(5):1578–83. http://dx.doi.org/10.1021/es801532f.
[102] Magdeldin M, Kohl T, Järvinen M. Process modeling, synthesis and thermodynamic evaluation of hydrogen production from hydrothermal processing of lipid extracted algae integrated with a downstream reformer conceptual plant. Biofuels 2016;7(2):97–116. http://dx.doi.org/10.1080/17597269.2015.1118785.
[103] Özdenkći K, De Blasio C, Sarwar G, Melin K, Koskinen J, Alopaeus V. Techno-economic feasibility of supercritical water gasification of black liquor. Energy 2019;189:116284. http://dx.doi.org/10.1016/j.energy.2019.116284.
[104] Ribeiro AM, Grande CA, Lopes FV, Loureiro JM, Rodrigues A. A parametric study of layered bed PSA for hydrogen purification. Chem Eng Sci 2008;63(21):5258–73. http://dx.doi.org/10.1016/j.ces.2008.07.017.
[105] Andersson E, Harvey S. Comparison of pulp-mill-integrated hydrogen production from gasified black liquor with stand-alone production from gasified biomass. In: International Conference on Efficiency, Cost, Optimization, Simulation, and Environmental Impact of Energy Systems. Energy 2007;32(4):399–405. http://dx.doi.org/10.1016/j.energy.2006.06.021.
[106] Magdeldin M, Kohl T, De Blasio C, Järvinen M, Won Park S, Giudici R. The BioSCWG project: Understanding the trade-offs in the process and thermal design of hydrogen and synthetic natural gas production. Energies 2016;9(10):838. http://dx.doi.org/10.3390/en9100838.
[107] Heyne S, Harvey S. Impact of choice of CO2 separation technology on thermo-economic performance of Bio-SNG production processes. Int J Energy Res 2014;38(3):299–318. http://dx.doi.org/10.1002/er.3038.
[108] Peduzzi E. Biomass to liquids. (Ph.D. thesis), Switzerland: EPFL; 2015, URL https://infoscience.epfl.ch/record/204975.
[109] Tock L, Gassner M, Maréchal F. Thermochemical production of liquid fuels from biomass: Thermo-economic modeling, process design and process integration analysis. Biomass Bioenergy 2010;34(12):1838–54. http://dx.doi.org/10.1016/j.biombioe.2010.07.018.
[110] Albarelli J, Onorati S, Caliandro P, Peduzzi E, Maréchal F, Ensinas A. Thermo-economic optimisation of integrated ethanol and methanol production in the sugarcane industry. Chem Eng Trans 2014;39:1741–6. http://dx.doi.org/10.3303/CET1439291.
[111] Maréchal F, Kalitventzeff B. Process integration: Selection of the optimal utility system. In: European symposium on computer aided process engineering-8. Comput Chem Eng 1998;22:S149–56. http://dx.doi.org/10.1016/S0098-1354(98)00049-0.
[112] Gerber L, Fazlollahi S, Maréchal F. A systematic methodology for the enviro-nomic design and synthesis of energy systems combining process integration, life cycle assessment and industrial ecology. In: Selected papers from ESCAPE-22 (european symposium on computer aided process engineering - 22), 17-20 June 2012, London, UK. Comput Chem Eng 2013;59:2–16. http://dx.doi.org/10.1016/j.compchemeng.2013.05.025.
[113] Streamlit Inc. API reference — Streamlit 0.88.0 documentation. 2021, URL https://docs.streamlit.io/en/stable/api.html.
[114] McKinney W. Data structures for statistical computing in python. In: Proceedings of the 9th Python in science conference. 2010, p. 56–61. http://dx.doi.org/10.25080/Majora-92bf1922-00a.
[115] Sievert C. Interactive Web-Based Data Visualization with R, Plotly, and Shiny. 2019, URL https://plotly-r.com/.
[116] Fallon J. Topsis-jamesfallon: Implementation of TOPSIS decision making. 2016, URL https://gitlab.com/jamesfallon/topsis-python.
[117] Opricovic S, Tzeng G-H. Compromise solution by MCDM methods: A comparative analysis of VIKOR and TOPSIS. European J Oper Res 2004;156(2):445–55. http://dx.doi.org/10.1016/S0377-2217(03)00020-1.
[118] Glasser GJ, Winter RF. Critical values of the coefficient of rank correlation for testing the hypothesis of independence. Biometrika 1961;48(3/4):444–8. http://dx.doi.org/10.2307/2332767.