Check for updates

Biofpr
Biofuels Bioproducts & Biorefining

Superstructure optimization of a microalgal biorefinery design with life cycle assessment-based and economic objectives

Johannes Kopton, (iD) Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg, Germany; Otto-von-Guericke University Magdeburg, Magdeburg, Germany

Liisa K. Rihko-Struckmann (iD) 🐦 **, Laura König-Mattern,** (iD) Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg, Germany

Kai Sundmacher, (iD) Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg, Germany; Otto-von-Guericke University Magdeburg, Magdeburg, Germany

Abstract: We present a multi-objective optimization framework for the design of an algal biorefinery with multiple target products. Four environmental endpoint indicators and the economic performance are used as objective functions following the life cycle assessment (LCA) methodology. The process alternatives are modeled as a superstructure covering a total of 720 feasible routes. The proposed method optimizes not only the superstructure route but simultaeously discrete and continuous parameters within the process units. A multi-objective genetic algorithm (MOGA) is used to solve this nonlinear mixed-integer optimization problem to design the extraction procedures for all macromolecular fractions. For the extractions, liquid–liquid equilibria (LLE) are predicted with quantum chemical calculations. A microalgae biorefinery for the marine diatom *Phaeodactylum tricornutum* is considered as case study, including the cultivation and extraction-supported fractionation of the wet algal biomass to harvest the target products eicosapentaenoic acid (EPA), laminarin and fucoxanthin. 2-Butanol proved to be the preferred solvent for the initial extraction step of wet biomass. Nutrients and solvent production cause most of the environmental impact of the overall process. A Python package for integrating LCA with multi-objective superstructure optimization is provided as open source software. It is applicable to any process system design task involving environmental objectives. © 2023 The Authors. Biofuels, Bioproducts and Biorefining published by Society of Industrial Chemistry and John Wiley & Sons Ltd.

Supporting information may be found in the online version of this article.

Correspondence to: Liisa K. Rihko-Struckmann, Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg, Germany. E-mail: rihko@mpi-magdeburg.mpg.de

SCI
where science meets business

Introduction

In recent decades, microalgae have attracted attention as a sustainable biomass resource. In contrast to crops, microalgae cultivation does not require arable land. Sustainability benefits include avoidance of land competition with food crops and lower demand for fresh water. Microalgae cells contain varying amounts of carbohydrates, lipids and proteins, including molecules of high nutritional or pharmaceutical value such as Ω-3-fatty acids, polysaccharides or pigments. Following a biorefinery approach, in this work, we consider the production of multiple marketable products from the same feedstock for full valorization of the biomass. However, the optimal design of a microalgal biorefinery is challenging. The sequential processing steps, starting with the cultivation, up to the separation of the target products, are highly interdependent and each processing step could be performed by several process units. Furthermore, economic and environmental interests commonly compete. In this study we therefore present a framework for global superstructure optimization. A biorefinery producing algae-based products might cause significant environmental impact during cultivation, harvesting and downstream processing, since nutrients, flocculants, solvents, heat and electrical energy are needed. To evaluate and compare overall environmental sustainability, life cycle assessment (LCA) enjoys widespread popularity. In order to not only analyze but also improve process sustainability for emerging technologies, the dynamically generated LCA results can be used as objective functions for life cycle optimization (LCO). We present here an optimized biorefinery, which produces and fractionates the wet algal biomass of *Phaeodactylum tricornutum*.

Sustainability analysis and LCA modeling

The life cycle assessment is an established method to estimate environmental impacts by a production system. The ISO standards 140 40 and 140 44 define the framework to carry out an LCA. The standards outline the overall approach and the iterative stages of the analysis: goal and scope definition; life cycle inventory (LCI, data collection); life cycle impact assessment (LCIA) and interpretation.

In this study, the system boundary for the LCA follows the *cradle-to-gate* principle: we consider the biomass cultivation, harvesting, extraction, separation, and purification (removal of solvents) of commercial high-value products inside the system boundary. Construction, maintenance and dismantling of production facilities are excluded. Since the marine diatom *Phaeodactylum tricornutum* contains several marketable substances, we analyze here the potential to produce multiple products. In this study, product compositions are not defined *a priori* but left for the optimizer. The functional unit is therefore defined as monetary value (EUR 1) of marketable products portfolio (one or more products together creating EUR 1 of revenue). In economic terms, this is called *total revenue*. The environmental impacts are reported in relation to the functional unit. Accordingly, we aim to identify environment-friendly product portfolios with promising economic returns.

Life cycle assessment model and impact analysis

The quantities of nutrients, solvents or additives, energy demand and disposal of materials are estimated based on literature data or obtained by process simulations. The LCA background data are taken from the Ecoinvent 3.7.1[1] database. Using the Ecoinvent data, all elementary flows related from or to the materials and energy supply for the whole supply chains was prepared applying the OpenLCA[2] software.

For the life cycle impact assessment (LCIA), we follow the European product environmental footprint (PEF) methodology,[3] which features 16 midpoint impact categories and use global normalization factors.[4] The PEF impact categories, the emission–damage pathways and resource–depletion models are given in Supporting Information, Table S1. We grouped the normalized impact category results into four endpoint environmental optimization objectives (climate change, human health, ecosystem quality, resources) using the PEF weighting factors to weight the impact categories within the environmental objectives.

Superstructure architecture and optimization objectives

Designing a complex separation flow scheme for an algal biorefinery is challenging. Traditionally, in an early-stage process conception, candidate routes are evaluated without precise mass and energy balances, having only rough estimations, for example for the molecular distribution, partition and energy demand. However, a more reliable analysis of the pathways is possible only with quantitative process flow sheet simulations relying on precise mass and energy balances, using thermodynamically consistent phase equilibrium models for separations. Furthermore, in an LCA, information regarding the exact emissions and composition of the side and waste streams are of particular importance.

However, they are rarely available during the preliminary process design stage, and they are obtained reliably by physical tail pipe measurements of existing units.

Superstructure optimization can be applied to process system design to identify the most favorable processing routes and operational decisions by maximizing the profit or other objectives. Gebreslassie *et al.*[5] and Gong and You[6] included sustainability metrics in the superstructure optimization for an algae-based biorefinery producing biofuel. However, these studies consider one economic and only one environmental goal, namely global warming potential (GWP). The optimization problems in Gebreslassie *et al.*[5] and Gong and You[6] were restricted to a closed mathematical form and formulated as a conventional mixed integer nonlinear programming (MINLP) problem with a two-dimensional optimization vector.

A superstructure model consists of two main components: processes and flows. In the language of graph theory, the processes are nodes, which are connected by edges representing the flows. For illustration, a schematic network graph is shown in Supporting Information, Fig. S1. The alternatives, for example selection of chemicals for one processing step, constitute a layer. Figure 1 illustrates the network graph naming the layers and all potential process units for the present biorefinery. To exemplify a layer, the flocculation methods using either $AlCl_3$ (C2), Al_2SO_4 (C3), chitosan (C4), or pH change (C5) establish *layer C*, also including the option to omit the use of any flocculation chemicals (C1, empty). The superstructure allows the complete crosslinking between the processes of the previous layer as well as of the next layer. The flows are modeled as vectors of partial flows constituting of *species*, which can be a

chemical substance (e.g. $NaSO_4$), biomacromolecular fraction (e.g. protein), energy form (e.g. electricity) or other forms of resources (e.g. land use). A complete list of the species used in this study is given in Supporting Information, Table S2.

A unique feature of our optimization framework is that it attempts to find the optimal superstructure route and *simultaneously* the most favorable in-process parameters (e.g., water/solvent ratios for extractions or operating temperatures). Furthermore, the framework allows the system to either have one or multiple outputs, splitting the superstructure into optional substructures (branches) leading to various products.

The life cycle impact assessment (LCIA) results are linked with the process design by optimizing four environmental goals plus the profit as a five-dimensional objective function vector. The optimization problem is formulated as follows:

$$\min \vec{f}(\vec{x}) \qquad (1)$$

where

$$\vec{f}(\vec{x}) = \begin{bmatrix} f^{climate}(\vec{x}) \\ f^{health}(\vec{x}) \\ f^{ecosystem}(\vec{x}) \\ f^{resources}(\vec{x}) \\ f^{economy}(\vec{x}) \end{bmatrix} \text{ and } \vec{x} = \begin{bmatrix} \vec{x}^{choice} \\ \vec{x}^{param} \end{bmatrix} \qquad (2)$$

The objective function $\vec{f}$ consists of the four environmental goals: climate, health, ecosystem and resources. To exemplify, six impact categories contribute to
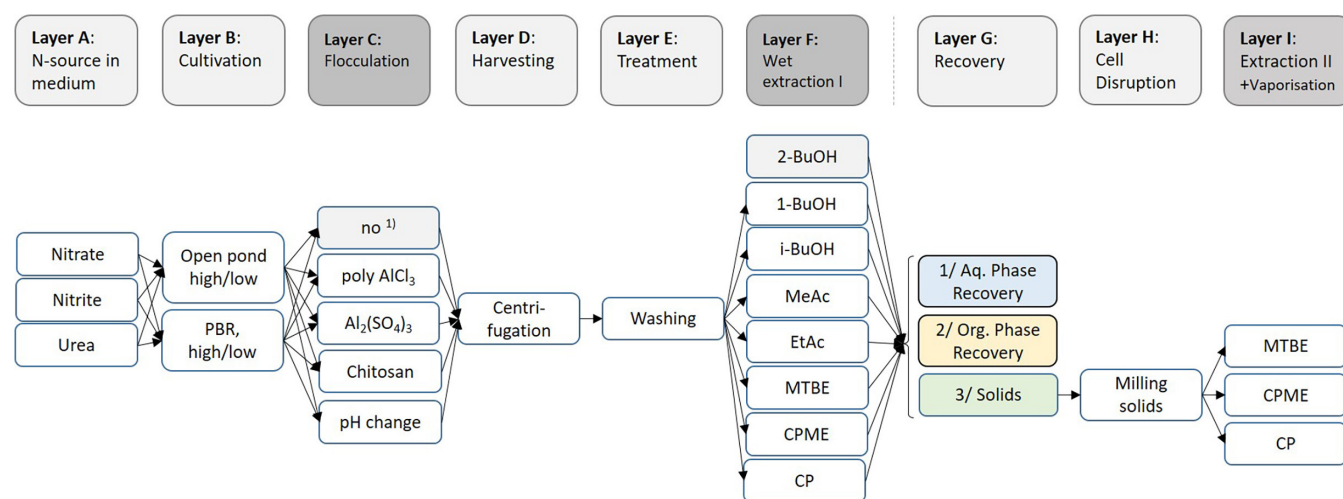

Figure 1. Superstructure network graph for the case study: Wet extraction biorefinery for *Phaeodactylum tricornutum*. Cultivation modus: high (open pond 0.06 wt%), low (open pond 0.03 wt%) biomass concentration. *i*-BuOH, *iso*butanol; MeAc, methyl acetate; EtAc, ethylacetate; MTBE, methyl tert-butyl ether; CPME, cyclopentyl methyl ether; CP, cyclopentane; PBR, photobioreactor.

the goal function *ecosystem*: (1) freshwater and (2) terrestrial acidification, (3) terrestrial, (4) freshwater, and (5) marine eutrophication as well as (6) the freshwater ecotoxicity. Supporting Information, Table S1 lists the impact categories contributing to each environmental goal. The optimization variable, vector $\vec{x}$, consists of the superstructure process choices, $\vec{x}^{\text{choice}}$, and the operational parameters in the process units, $\vec{x}^{\text{param}}$.

As there are five objective functions, we obtain a set of Pareto-optimal solutions – where one Pareto-optimal solution is defined as a solution in which no objective can be further improved without impairing another one. The final selection of a recommended design depends on the relative weighting of the objective functions, which is a human value decision. After the optimization, this decision can be taken with full knowledge of the entire Pareto set following an *a posteriori* decision-making approach.[7]

Algorithm selection

As a conventional MINLP solver aims to find the exact global optimum, the respective optimization problems need to adhere to a strict mathematical form, e.g. (piecewise) continuity, differentiability, and quasi-convexity. However, the objective function vector $\vec{f}(\vec{x})$ in our study contains various nonalgebraic operations (e.g. the solution of an embedded linear sum assignment problem), leading to a noncontinuous, nonconvex optimization problem. We therefore apply a global population-based metaheuristic.[8] This allows the approximate solution to be found without requirements to the model other than pointwise evaluation. For such problems, multi-objective genetic algorithms are highly effective.[9] We selected the well-established *nondominated sorting genetic algorithm 2* (NSGA-II).[10]

The superstructure framework is implemented in Python using the open source framework for multi-objective optimization *Pymoo*. The *Pymoo* Python library provides an efficient implementation of NSGA-II. It is highly flexible, allowing us to implement custom sampling, crossover, and mutation methods to handle our mixed-integer search space.

Sampling, mutation, and crossover

As a genetic algorithm, NSGA-II is inspired by biological evolution to approximate the optimal results. Initially, a population of random solutions (individuals) is generated (sampling). In this context, the optimization variable $\vec{x}$ is considered a genotype, while the objective function $\vec{f}(\vec{x})$ represents the phenotype. In each iteration (generation), the genotype is randomly modified (mutation) and genotypes of different individuals are combined (crossover). In the next step, the phenotypes are compared and the best (in the sense of *nondominated*) individuals form the population of the next generation.

The start population is sampled using uniform distributions within the respective variable bounds $x_i \sim U(x_i^{\min}, x_i^{\max})$ and a population size of 201 individuals. For mutation, different methods need to be used for the different variable types: Integer variables are increased or decreased by 1 with a probability of p_{integer} each (respecting the variable bounds). Boolean variables are flipped with a probability of p_{boolean}. Continuous variables are mutated by adding a normally distributed random variable $Z \sim N(0, \sigma^2)$ centered on zero with a standard deviation of σ (again respecting the variable bounds).

In the case study, the algorithm converges reliably with the hyperparameters $p_{\text{integer}} = p_{\text{boolean}} = \sigma = 0.2$, which are chosen as default values for the framework.

Economical metrics and product value estimation

One important objective in this study is to achieve the highest possible operating margin (OM), which can be defined in terms of operating cost (OC) and total revenue (TR): $\text{OM} = 1 - \text{OC}/\text{TR}$. In order to maximize operating margin, OC/TR is minimized and constitutes the fifth optimization objective $f_5(\vec{x})$. The expected selling prices of the products are estimated using wholesale offers as guidelines. The market value of natural substances can be modeled as a function of purity and for the case study, they were interpolated linearly from available wholesale offers (see Supporting Information, Table S3). In this framework, each product flow is modeled as being marketed for the target substance, which respectively yields the highest selling price, considering the calculated process-dependent purity. Other valuable substances that might be present in the same product flow are not included in the pricing estimation but are considered instead as (nontoxic) impurities.

The operating costs were calculated as shown in Eqn 3, with $\dot{F}$ representing a vector of mass and energy flows into (makeup) or out of (emission) the production system and p representing purchasing/disposal prices per species. In the case study, disposal costs were assumed to be negligible, $p^{\text{disposal}} = 0$.

$$\text{OC} = \left\langle p^{\text{purchase}}, \dot{F}^{\text{makeup}} \right\rangle + \left\langle p^{\text{disposal}}, \dot{F}^{\text{emission}} \right\rangle \quad (3)$$

Integration of LCA results

Similar to the economic costs, environmental impacts are calculated as a linear function of mass and energy flows (Eqn 4). However, here the coefficients *A* are matrices with the species (i.e. chemicals and energy forms) as columns and LCIA categories (e.g. global warming potential) as rows. This yields a vector $\dot{L}$ of LCIA results. Here, impacts from emissions were included in the analysis.

$$\dot{L} = A^{\mathrm{production}} \cdot \dot{F}^{\mathrm{makeup}} + A^{\mathrm{emission}} \cdot \dot{F}^{\mathrm{emission}} \qquad (4)$$

In our Python package, Eqns 3 and 4 are implemented with one additional dimension each to account for the contributions of the individual processing units. However, the sectoral contributions are relevant only for post-hoc analysis and do not influence optimization, so for the optimization we simplified the mathematical execution.

The individual LCIA results, structured as coefficient matrices *A* were extracted from the Ecoinvent database using the openLCA program over the IPC interface.[11] As LCIA is a linear operation, the calculation of LCIA results for all relevant candidate species prior to the overall life cycle inventory (LCI) is possible. This improves calculation performance significantly without changing the results. It allows fast online assessment of the LCA during optimization. The details for the calculation procedure are given in the supporting information.

For each environmental goal (climate, health, ecosystem, and resources), the respective impact categories are then aggregated ($\widetilde{L}_i$) and related to the functional unit, namely the total revenue (Eqn 5). This constitutes all the environmental objective functions:

$$\vec{f}_i(\vec{x}) = \frac{\widetilde{L}_i(\vec{x})}{\mathrm{TR}(\vec{x})} \qquad (5)$$

Biorefinery optimization

With high industrial relevance in the marine bioeconomy, we optimize a biorefinery valorizing wet algal biomass of the diatom *P. tricornutum*. However, the methodological framework we presented above, is not limited to this case study but is easily applicable to other comparable separation processes in the chemical or biotechnical industry.

Biorefinery for wet algal biomass

Depending on the strain, *P. tricornutum* cells contain varying amounts of carbohydrates, lipids, and proteins, including specific molecules with extreme high nutritional

or pharmaceutical values such as polyunsaturated fatty acids, polysaccharides, or pigments. Conventionally, the algal biomass is dried, and the utilization is focused on a single target molecule, while residual algal biomass is used in low value applications, e.g. as animal feed or for anaerobic digestion.[12]

In our biorefinery approach, the biomacromolecules (e.g. lipids, carbohydrates and proteins) are fractionated, and multiple marketable products are utilized. This way, the algal biomass is maximally exploited, which leads to optimal economic and environmental efficiency. *P. tricornutum* has several marketable high-value products: the pigment fucoxanthin, polyunsaturated eicosapentaenoic fatty acid (EPA) and laminarin, a β-glucan polysaccharide with health-supporting properties.[13]

For the cultivation of *P. tricornutum* in pasteurized sea water, we consider three alternative N-sources ($NaNO_3$, $NaNO_2$, and urea).[14] Seawater is one important makeup flow into the system. After insertion into the system, it is considered a mixture of water and NaCl. Experimental results from open pond (OP) or photobioreactor (PBR) cultivation under Mediterranean climate conditions[14] are used to estimate the growth rate of the algal biomass. The composition of fatty acids and their polarity originated from an earlier study.[15] Four optional cultivation conditions were assessed, and the corresponding area requirements and productivities are shown in Table 1.

The biomass composition (ash, carbohydrate, protein and lipid amount) and the concentration of marketable products (laminarin, fucoxanthin and EPA) depend on the applied N source (*layer A* in Fig. 1), reactor type (OP or PBR in *layer B*) and the growth modus (A = high or B = low in Table 1). All

Table 1. Land demand, electric energy demand and biomass growth productivity for four cultivation options. OP = open pond, PBR = photobioreactor. Growth modus: *A* high and *B* low. Subscripts DW = dry weight, med = culture medium.				
Property	OP *A*	OP *B*	PBR *A*	PBR *B*
Obtained BM conc. Wt%[a]	0.06	0.03	0.1	0.06
Culture mass (kg$_{med}$ m^{-2}$_{land}$)	100[d]	100[d]	51.5[d]	51.5[d]
Electricity kJ kg^{-1}$_{med}$	32[b]	32[b]	95.5[c]	95.5[c]
Productivity g$_{DW}$ kg^{-1}$_{med}$ d^{-1}[a]	0.14	0.14	0.035	0.035
[a]Yang *et al.*[15] [b]Sarat Chandra *et al.*[16] [c]Porcelli *et al.*[17] [d]For direct land use for LCA, data from Silva Benavides *et al.*[14]				

correlating equations for the biomass composition depending on medium and cultivation conditions are obtained from the literature (summarized in Supporting Information, Table S4). To concentrate the cultivation broth, a flocculant (*layer C*) can be selected. The dynamics of algal biomass concentration are specific to the flocculant,[18] and the flocculation and centrifugation parameters are listed in Supporting Information, Table S5. The wet algal biomass is separated from the aqueous cultivation medium by centrifugation (*layer D*), and the biomass is concentrated up to a level of 20% dry weight (DW).

Recycling of the cultivation medium is essential in a biorefinery, contributing significantly to the economic and environmental sustainability. Recycling reduces the demand of makeup chemicals in addition to lowering the amount of sewage produced.

However, the culture medium recycling ratio is not unlimited as the used culture medium gets contaminated by extracellular metabolites. We therefore included the recyclability factor $r_{max} \in [1, 0]$.

In the optimization, the recycling poses a remarkable challenge, because some disposal flows could possibly be recycled and a best assignment of disposals needs to be found. The economic value of the recycling stream was taken as decisive factor for the recycling optimization. The mathematical approach to solve the arising linear sum of assignment problem (LSAP) to optimize the recycling is given in the supporting information.

The biorefinery flowsheet is illustrated schematically in Fig. 2. The proposed process system separates lipids, pigments, proteins, and carbohydrates from wet biomass, omitting the energy intensive drying. The first extraction (unit Extr-I in Fig. 2) is carried out for undisrupted, wet cells. As described in detail in our previous study,[19] the water/solvent ratio in Extr-I is adjusted initially to establish only one liquid phase. After a desired extraction time, subsequently the water/organic solvent ratio is decreased by inserting additional organic solvent into Extr-I to establish a separation into aqueous and organic liquid phases. Following the prediction as elaborated in an earlier work,[19] the two-liquid phase system is expected to form an autonomous partition of carbohydrates and polar molecules. Based on computational predictions,[19] it could be expected that solid residual biomass is left over after Extr-I.

For the solid residual biomass separated by filtration, bead milling was selected as most suitable cell disruption method assuming an electrical energy demand of 0.14 kWh kg^{-1}.[20] After cell disruption, the residual solid biomass is conducted to the second extraction unit, Extr-II. Here, predominantly neutral lipids, i.e. triglycerides and fucoxanthin are separated from remaining carbohydrates using the optimizer-selected solvent (superstructure selection in *layer I*). After
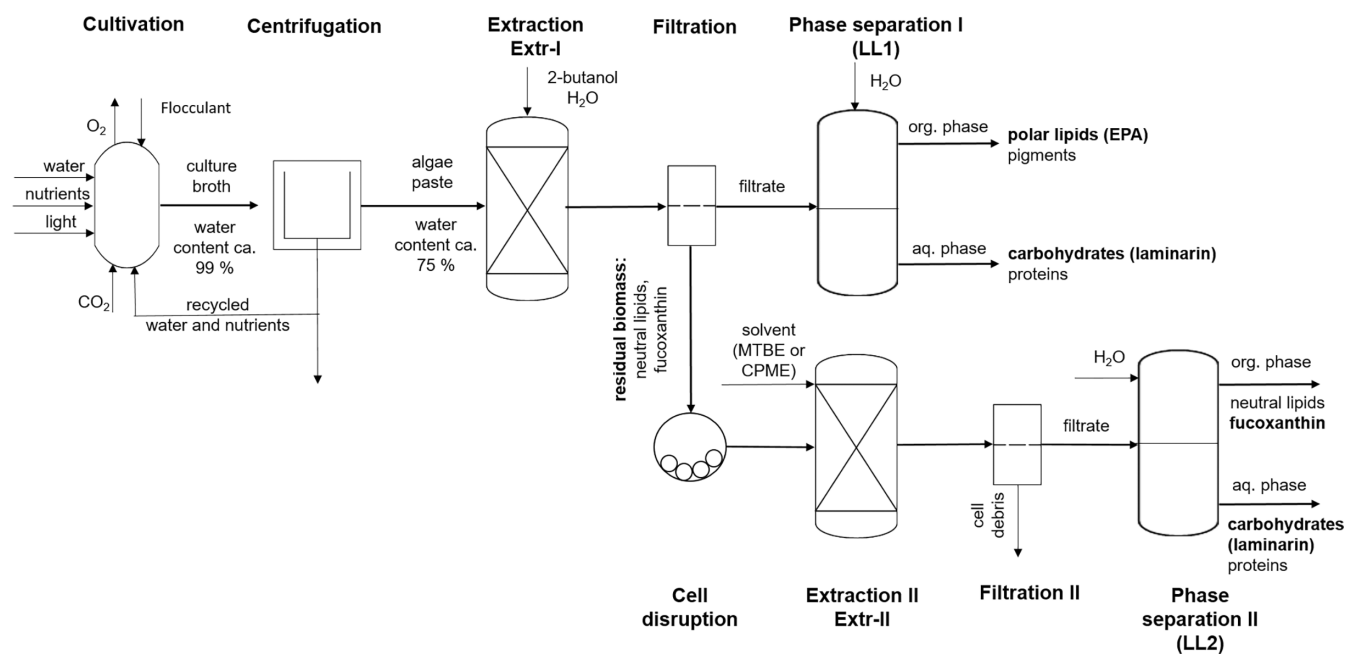

Figure 2. The conceptual process design to separate marketable fractions of the marine diatom *Phaeodactylum tricornutum*. The superstructure is optimized in order to identify the economically most profitable product portfolio with simultaneous selection of the most environmentally benign solvents for extractions.

the extractions, the products are recovered by vacuum evaporation and the organic solvents are recycled.

Following the first extraction, Extr-I (*layer F*), the water and organic solvent soluble molecules are separated and dried (branches G1-I1 and G2-I2, correspondingly). The solid residual biomass is utilized following one of the paths/branches G3-H3-I3, G3-H3-I4 or G3-H3-I3. In a separate optimization run, the optimizer was allowed to omit the branches (valorization of solid biomass) completely, leading possibly to less revenue but in the same time to lower costs and environmental impact.

Estimation of solubility, phase separation, partition coefficients, and solvent recovery

In the computational screening,[19] eight solvent candidates are identified for unit Extr-I and three for Extr-II. The solvent selection is challenging in the early stage process design, because a process flow sheet simulation requires the knowledge of thermodynamic properties to simulate the liquid–liquid equilibrium (LLE) for the solvent/water mixtures and the partition of the exploitable components between the phases. The experimental LLE information, or reliable interaction parameters for mutual solubility estimation were lacking for a few solvent/water mixtures. We predicted the binary solvent/water phase equilibria computationally following the quantum chemical approach (COSMO-RS, conductor like screening model for real solvents) similar to an earlier publication.[19] The solubility and the partition coefficients of the biomass constituents (neutral and polar lipids, fucoxanthin, protein, and carbohydrates) in the respective solvent-mixtures were predicted using the COSMO-RS and reference molecules were used as model compounds for protein and carbohydrate fractions. The COSRMO-RS estimated solubilities of the biomass constituents are given in Supporting Information, Table S6.

Solvent evaporation was simulated using the simulation tool DWSIM[21] to estimate the solvent specific energy demand. The thermodynamic compound data and the UNIQUAC/UNIFAC-parameters (for cyclopentyl methyl ether, CPME) for the solvents were taken from *ChemSep* data base[22] to estimate the activity coefficients for vapor–liquid equilibrium (VLE). For the heat sensitive fucoxanthin, we assumed a maximum temperature tolerance of 37 °C (requiring subatmospheric pressure). The solvent use contributes significantly to the costs and environmental impact, therefore we take the solvent-to-biomass ratio as subject to optimization. We assumed that a maximum of 75% of the solvent amount was recyclable.

The Ecoinvent database did not contain the life cycle impact information for all the candidate solvents. To estimate the environmental interventions, therefore, a few assumptions were made: for cyclopentane (CP), the LCI data of cyclohexane, and for CPME, the LCI data for methyl cyclopentane were taken as substitutes. For the land use impact category, *industrial area* was selected, and the Italian grid mix was assumed as electric energy supply, corresponding to the selected Mediterranean cultivation conditions.

Results and discussion

The superstructure consists of 720 unique paths, while the Boolean and continuous parameters in the process units create a multidimensional search space for each path. To visualize the high number of feasible solutions, Fig. 3 shows 500 randomly sampled feasible solutions as a pairwise scatter plot. The majority of generated solutions are unanimously suboptimal, i.e. configurations are far from optimal in all impact goals at the same time. For example, excessive solvent use is not only expensive (economy category, expressed as high cost/revenue ratio) but it is climate-damaging at the same time. As Fig. 3 shows, the sampled solutions therefore form linear-resembling correlations. Furthermore, the economic objective shows a calculated cost/revenue value up to 300 for a nonoptimized point (a randomly sampled solution). Most of the points are unanimously suboptimal. Such a point is mathematically feasible (allowed), but operationally very far from economically feasible. This clearly demonstrates the need for the optimization presented below. To note, the Pareto-optimal points are located in the immediate vicinity of the origin in Fig. 3.

Using the NSGA-II algorithm, we mapped out the Pareto-optimal frontier with 201 solutions, covering it as uniformly as possible using the crowding distance metric.[10] Clusters of optimal points can be observed resulting from the discrete variables (e.g. superstructure path or a selected solvent), where the variation of continuous variables is responsible for the extent of the clusters, see Fig. 4.

Optimal superstructure path

All optimal results follow the same superstructure path leading to several marketable products: aqueous and organic phase from both extraction units, Extr-I and Extr-II respectively (see Fig. 5).

In *layer A*, an inexpensive N source, urea, was the optimal selection, because our model predicts this leading to high concentrations of EPA. This is in line with the experimental findings of Perez-Lopez *et al.* (2014).[23] They reported that the use of urea as an N source causes clearly lower
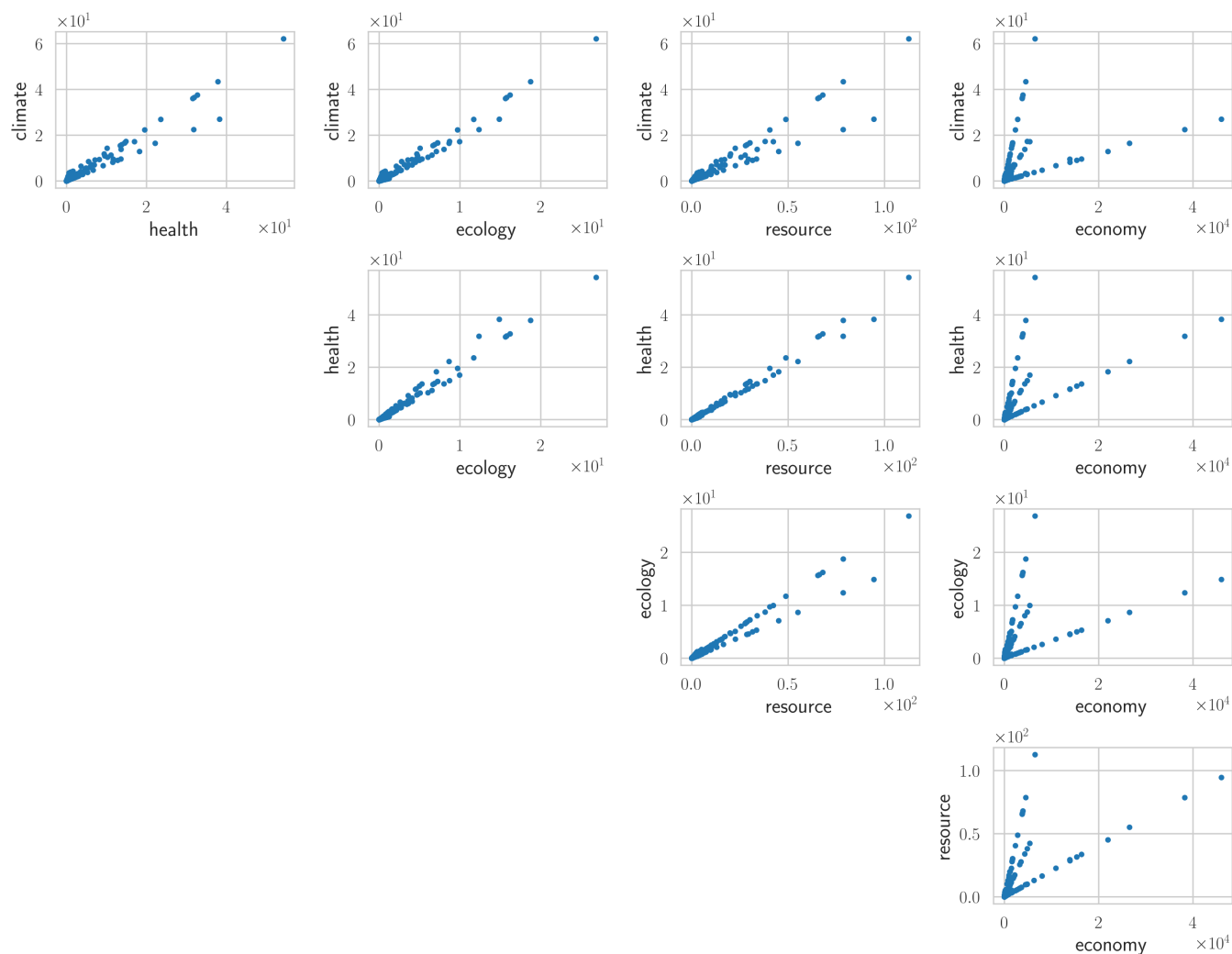
Figure 3. Five hundred randomly sampled feasible solutions in the five-dimensional objective space presented as projections in two-space combinations of all objective functions. All objective functions (environmental impacts and economics expressed as production cost/revenue) are to be minimized, and the best solutions approach the origin.

environmental burdens in acidification and eutrophication than NaNO3. In *layer B*, open ponds are preferred over closed photobioreactors despite larger land use and lower biomass productivity, due to the lower electrical energy demand,[24,25] compared to photobioreactors. Growth modus A – leading to a high biomass concentration (see Table 1) – is the preferred cultivation modus here. In *layer C*, omission of flocculant chemicals was found to be optimal. The use of flocculation would clearly reduce the cost for centrifugation (*layer D*). However, any of the flocculation chemicals increase heavily the environmental burdens as they leave the process mainly as disposed culture medium (emissions). Furthermore, the use of flocculants limits the recyclability of the culture medium, leading to higher overall nutrient costs and emissions. This conclusion agrees fully with our previous findings.[18,26] In *layers D* and *E*, the centrifugation

and biomass washing are mandatory processing steps without alternatives in our model. Centrifugation is necessary to remove the excess water, and washing to lower the flocculants or nutrient fraction in algal biomass.

For the extraction step performed on wet biomass, Extr-I corresponding to *layer F*, out of seven solvent candidates, 2-butanol was identified as the most favorable in all optimal routes. First, the high water amount in the miscibility range of 2-butanol/water solution is favorable because it lowers organic solvent demand, and enables solvent/water recycle to a high extent. Second, according to the COSMO-RS predictions, 2-butanol preferentially solves the polar lipids rather than the neutral lipids and pigments, which remain in the solid residue. Similarly, proteins and carbohydrates are soluble in the aqueous 2-butanol mixture, and subsequently they are recoverable in the aqueous phase

 | *Biofuels, Bioprod. Bioref*. **17**:1515–1527 (2023); DOI: 10.1002/bbb.2540
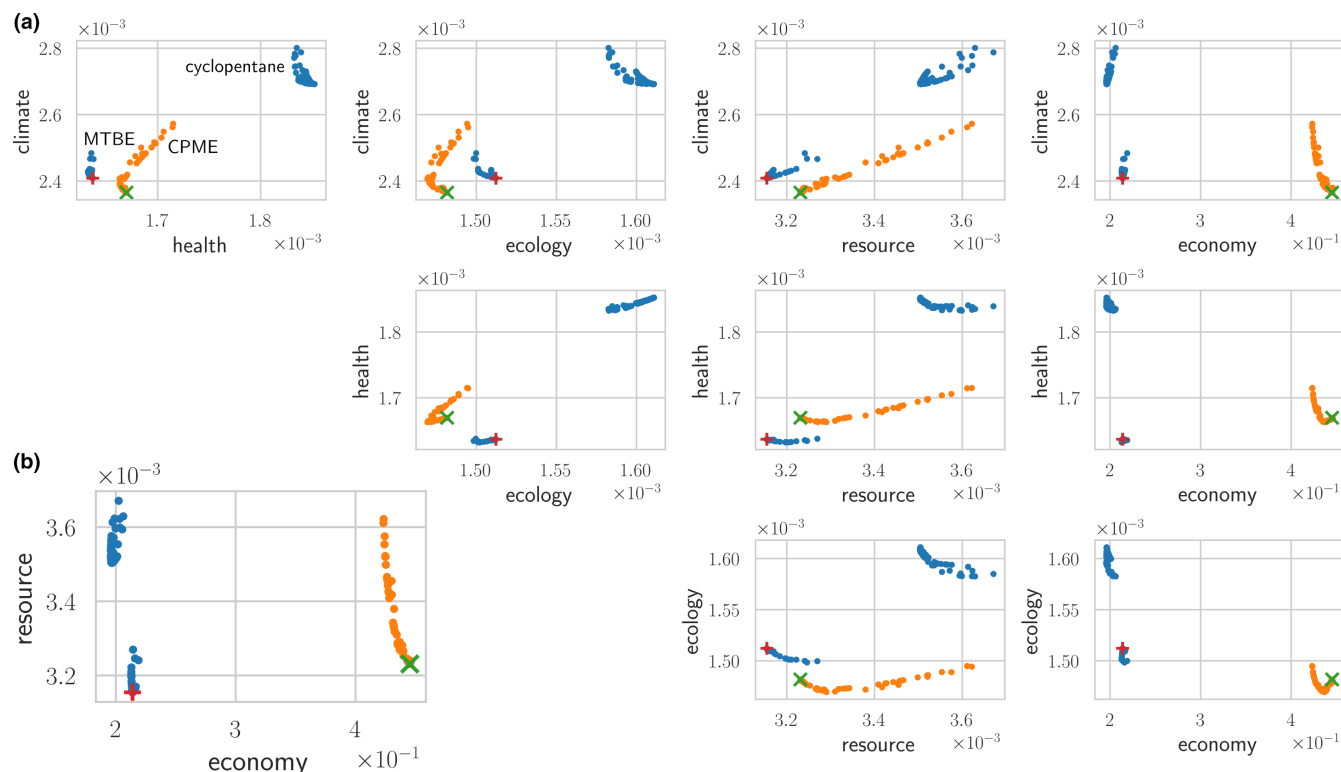
Figure 4. The Pareto frontier in objective space, (a) projected on two-dimensional spaces spanned by each combination of two different objective functions; (b) enlarged illustration for the projection space: economy – resource. Three distinct groups related to the solvent selection for the extraction of the residue solid. Example solutions are marked with a green cross (CPME, x) or red plus (MTBE, +).
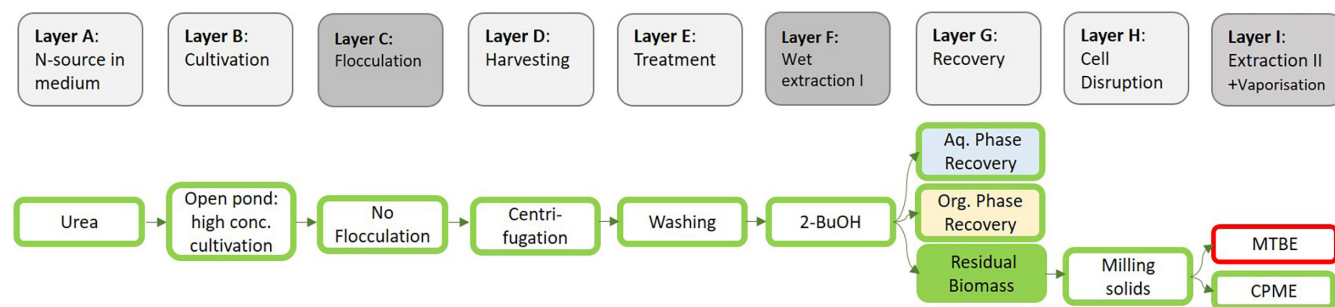

Figure 5. Optimizer-selected process units and solvents in the superstructure for the exemplary Pareto optimal solutions corresponding to x and + in Fig. 4. All Pareto-frontier points follow the same path until layer H: In layer I the solvent for Extr-II is MTBE (+) or CMPE (x). All branches on the recovery (layer G) are present in the exemplary solutions (no branch cut-off).

after the 2-butanol – water phase separation. Coincidentally, 2-butanol was identified previously as the most promising solvent for microalgae wet extraction by a computational screening.[19] For the second extraction, Extr-II, we used the same methodology to screen for solvents with high solubility of neutral lipids. All solvents with a predicted neutral lipid solubility equal or higher to that of hexane were considered.

Cyclopentyl methyl ether and methyl *tert*-butyl ether (MTBE) were identified as promising solvents.

Layers G (solvent recovery) and *H* (cell disruption) are unavoidable; the solvent recovery is mandatory to separate the products from solvents. To recover the products from the solid fraction; similarly, the cell disruption is compulsory. For layer *I*, the extraction of the solid biomass (Extr-II) and

subsequent removal of the solvent, we can report various optimal decisions depending on our valuation of the objectives. Two exemplary optimal solutions are highlighted with + and x in Fig. 4, which differ in the applied solvent for Extr-II. CPME (x) and MTBE (+) are the solvents in these exemplary solutions in Fig. 4. Cyclopentyl methyl ether leads to the lowest impact regarding the ecosystem and climate, due to low environmental burdens during the production of the solvent. However, CPME is significantly more expensive than MTBE and cyclopentane, which makes a remarkable difference in the economic outcome. Here, the revenue ratio would be lower, i.e. 0.2139 instead of 0.4452 when CPME is substituted by MTBE as solvent in Extr-II (*layer I*), respectively. At the same time, all the other criteria show only minor dependence on the solvent selection – for example, for the climate objective, the numeric value changes from 0.0023 to 0.0024 when CPME is substituted by MTBE. The economic dependence on the solvent choice is mainly due to the differences in the purchasing costs.

The cluster with the lowest operating costs (using solvent cyclopentane in the Extr-II, see Fig. 4) is associated with less optimal values in all four environmental impact category goals.

Optimal process model parameters

Parallel to the identification of the superstructure path, a number of discrete and continuous parameters in the models describing the processing unit are optimized. In cultivation (*layer B*) the cultivation conditions leading to high biomass concentration are preferred as they led to a high overall lipid content. Cultivation achieving low biomass concentration leads to a larger protein fraction, which was less preferred in the optimization. The Ω-3 fatty acid EPA is a very valuable biorefinery product available in the lipid fraction, which explains the superstructure selection to prefer high biomass concentration with high overall lipid content. Natural pigments and lipids are the compounds with

the most influence on the overall revenue from microalgal biomass.[12] The proteins in *P. tricornutum* are safe for human consumption, but they do not have special pharmaceutical or nutritional value. The proteins with their preserved techno-functionality are the most challenging and costly products to extract, and therefore they do not play a role in *P. tricornutum* biomass valorization.[12] The water-soluble proteins could be utilized as food (EUR 5 kg^{-1}), the water-insoluble proteins as animal feed (EUR 0.75 kg^{-1}).[27] In this study, however, the potential revenue from selling the protein fraction was considered negligible.

In *layers F* and *I*, the amount of solvent is a continuous variable. The resulting amounts of 2-butanol/water mixture range from 25.1 kg to 38.2 kg kg$^{-1}_{DM}$ in the Pareto-set depending on the weighting of the objective functions. It is important to note that 25.1 kg is the lower bound for the solvent amount set in optimization, so the Pareto frontier is touching this bound. The solvent/biomass (20% dry weight) mass ratio was constraint, because lower mass ratios than that would induce an immediate phase separation to aqueous and organic phases, which is undesired initially in Extr-I. However, the solvent amount could possibly be further lowered if an additional dewatering process is designed for the water removal. In the Extr-II unit, comparable to Extr-I, the optimal amounts of solvent/water mixtures was limited by the pre-defined lower bound. The obtained Pareto-optimal solvent amounts in Extr-II were 78.8 kg (MTBE), 59.8 kg (CPME) or 112.0 kg (cyclopentane) per kg$_{DM}$.

Table 2 summarizes the marketable products in the three branches. The purity of the products was freely optimized, and expected revenue of marketable products was estimated based on their purity. The product quantities and qualities in Table 2 correspond to one of the optimal solutions (x in Fig. 4). The most valuable product in the present biorefinery is laminarin, which is available as the aqueous product of LL1 (shown in Fig. 2) after the aqueous phase vaporization. The estimated yield of laminarin in this product stream is predicted to be 81.9%. The optimized purity of the β-glucan

Table 2. Available and marketable products, their purity and the predicted selling prices.

Source	Available products	Marketable product	Estimated purity/%	Predicted revenue EUR kg^{-1}
Aqueous phase I	Laminarin Carbohydrates	Laminarin	33.0	40.0
Organic phase I	Polar EPA, pigments	Polar EPA	20.8	1.99
Solid residue: Organic phase II	Lipids neutral, fucoxanthin	Fucoxanthin	7.3	1.68
Solid residue: Aqueous phase II	Proteins, carbohydrates, laminarin	-[a]	-	-

[a]Not marketable as laminarin, EPA or fucoxanthin; other methods of valorization are possibly marketable.

is 33.0 wt%, which corresponds to a predicted price of EUR 39.97 kg^{-1}.

After vaporization of the organic phase (mostly 2-butanol), the polar-bound portion of the fatty acid EPA is the second marketable product from LL1. The EPA purity of 20.8 wt% with a selling price of EUR 1.99 kg^{-1} is predicted. A minor proportion of fucoxanthin is present in this fraction but it is considered neither marketable nor value adding in this context.

After the optimization we can conclude that it is profitable to utilize the residual solid biomass. After the extraction with CPME, neutral lipids and fucoxanthin are present in the organic phase obtained from the unit LL2. Fucoxanthin, with an expected purity of 7.34% and a predicted price of EUR 1.68 kg^{-1}, is the valuable, marketable product from this stream. The aqueous phase of LL2 contains very low concentrations of the defined marketable products, which makes the further purification and valorization less promising. The aqueous phase after LL2 therefore does not contribute to the revenue.

Figure 6 illustrates the environmental impacts of one Pareto-optimal solution expressed for the 16 predefined categories constituting the PEF methodology. The values are normalized by global averages and the contributions of the process units are differentiated by colors. The highest relative impact is created in the categories *fossil*

resource use and *carcinogenic effects on human health*. The high overall energy demand during the production of the nutrients and solvents, and the electrical energy used during cultivation, contribute significantly to the impact. The freshwater eutrophication impact is mainly caused by the disposal of used culture medium (shown as impact by the centrifugation unit in Fig. 6). Significant improvements would be achievable by increasing the recycling ratio of nutrients and solvents. However, tolerable recycling ratios of solvents or cultivation media is crucial, depending on the algae species, and reliable empirical data are lacking on maximal feasible recycling rates in the present processes. Furthermore, in a real process, the release of waste nutrients to environment is mitigated by careful sewage treatment.

Several categories (e.g. climate change, acidification or noncarcinogenic effects on human health) show correlating differentiated contributions, which is explained by the remarkable contribution of energy demand on these impact categories. The energy demand influences significantly all four environmental impact goals as well as economic performance. Many decisions therefore do not evoke tradeoffs, and this leads to the relatively small Pareto frontier as displayed in Fig. 4. Furthermore, the nutrient (urea) production for the algae cultivation is a significant contributor in many of the categories as shown in Fig. 6.
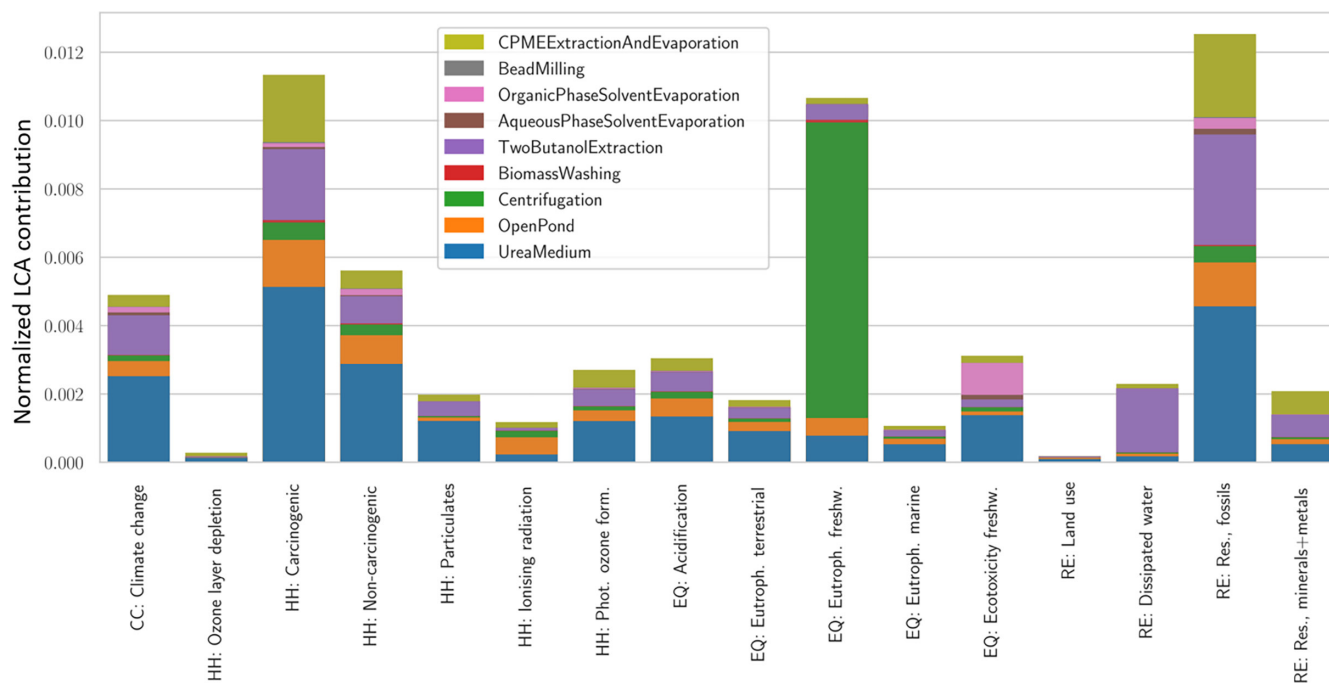

Figure 6. Normalized impact contributions (LCIA) for an exemplary optimal solution (x in Fig. 4) for 16 categories following the PEF methodology. Colors indicate the sectoral contributions. Impact groups (Fig. 4): CC, climate change; HH, human health; EQ, ecosystem quality; RE, resources. Global normalization factors and the PEF weighting factors used in the grouping.

Conclusions

A multi-objective superstructure optimization was applied successfully to design an environmentally friendly biorefinery for wet algal biomass by minimizing environmental impacts (LCA) and maximizing the economic operating margin. The multi-objective genetic algorithm NSGA-II found the Pareto-optimal configurations efficiently. Both superstructure route decisions and continuous and discrete parameters were optimized simultaneously, as demonstrated here for a biorefinery case study for marine diatom *P. tricornutum*. For wet algal biomass, the initial extraction with 2-butanol is a highly promising approach, because it enables to separate laminarin and the Ω-3 fatty acid EPA bound in polar cell structures. From the solid residue, fucoxanthin is recoverable as a marketable product. The high energy demand required for the multi-step separation procedures was the main contributor for the environmental impacts. Furthermore, nutrient and solvent selection, recycle ratios of cultivation medium and the recovery of the solvents increased significantly various environmental impacts.

The open source Python package developed for this work can easily be used for other process design tasks involving sustainability objectives. The framework is independently applicable for general superstructure optimization problems, optionally including the optimization of environmental impacts via life cycle optimization.

It is available at: https://git.mpi-magdeburg.mpg.de/pse-group/structimize.

Acknowledgements

References

1. Wernet G, Bauer C, Steubing B, Reinhard J, Moreno-Ruiz E and Weidema B, The ecoinvent database version 3 (part I): overview and methodology. *Int J Life Cycle Assess* **21**(9):1218–1230 (2016).
2. Ciroth A, ICT for environment in life cycle applications openLCA-a new open source software for Life Cycle Assessment. *Int J Life Cycle Assess* **12**:209–210 (2007).
3. 2013/179/EU: Commission Recommendation of 9 April 2013 on the use of common methods to measure and communicate the life cycle environmental performance of products and organisations Text with EEA relevance (2013).
4. European Commission, Joint Research C. global normalisation factors for the environmental footprint and Life Cycle Assessment LU: Publications Office (2017).
5. Gebreslassie BH, Waymire R and You F, Sustainable design and synthesis of algae-based biorefinery for simultaneous hydrocarbon biofuel production and carbon sequestration. *AIChE J* **59**(5):1599–1621 (2013).
6. Gong J and You F, Optimal design and synthesis of algal biorefinery processes for biological carbon sequestration and utilization with zero direct greenhouse gas emissions: MINLP model and global optimization algorithm. *Ind Eng Chem Res* **53**(4):1563–1579 (2014).
7. Hwang CL and Masud ASM, *Multiple Objective Decision Making, Methods and Applications: A State-of-the-Art Survey*. Springer-Verlag, Berlin and New York, p. 351 (1979).
8. Maniezzo V, Boschetti MA and Stützle TG, *Matheuristics: Algorithms and Implementations*. Springer, Cham, p. 214 (2021).
9. Zolpakar NA, Lodhi SS, Pathak S and Sharma MA, Application of multi-objective genetic algorithm (MOGA) optimization in machining processes, in, ed. by Gupta K and Gupta MK. Springer International Publishing, Optimization of Manufacturing Processes. Springer Series in Advanced Manufacturing, Cham, pp. 185–199 (2020).
10. Deb K, Pratap A, Agarwal S and Meyarivan T, A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans Evol Comput* **6**(2):182–197 (2002).
11. Srocka M, Garcia Casas M and Rickert J, Olca-IPC.Py-a client to communicate with an openLCA IPC server GreenDelta.
12. Ruiz J, Olivieri G, de Vree J, Bosma R, Willems P, Reith JH *et al.*, Towards industrial products from microalgae. *Energ Environ Sci* **9**(10):3036–3043 (2016).
13. Stiefvatter L, Neumann U, Rings A, Frick K, Schmid-Staiger U and Bischoff SC, The microalgae Phaeodactylum tricornutum is well suited as a food with positive effects on the intestinal microbiota and the generation of SCFA: results from a pre-clinical study. *Nutrients* **14**(12):2504 (2022).
14. Silva Benavides AM, Torzillo G, Kopecký J and Masojídek J, Productivity and biochemical composition of Phaeodactylum tricornutum (Bacillariophyceae) cultures grown outdoors in tubular photobioreactors and open ponds. *Biomass Bioenergy* **54**:115–122 (2013).
15. Yang YH, Du L, Hosokawa M, Miyashita K, Kokubun Y, Arai H *et al.*, Fatty acid and lipid class composition of the microalga Phaeodactylum tricornutum. *J Oleo Sci* **66**(4):363–368 (2017).
16. Sarat Chandra T, Maneesh Kumar M, Mukherji S, Chauhan VS, Sarada R and Mudliar SN, Comparative life cycle assessment of microalgae-mediated CO2 capture in open raceway pond and airlift photobioreactor system. *Clean Techn Environ Policy* **20**(10):2357–2364 (2018).
17. Porcelli R, Dotto F, Pezzolesi L, Marazza D, Greggio N and Righi S, Comparative life cycle assessment of microalgae cultivation for non-energy purposes using different carbon dioxide sources. *Sci Total Environ* **721**:137714 (2020).
18. Pirwitz K, Flassig RJ, Rihko-Struckmann LK and Sundmacher K, Energy and operating cost assessment of competing harvesting methods for D. salina in a β-carotene production process. *Algal Research* **12**:161–169 (2015).
19. König-Mattern L, Linke S, Rihko-Struckmann L and Sundmacher K, Computer-aided solvent screening for the fractionation of wet microalgae biomass. *Green Chem* **23**(24):10014–10029 (2021).

20. Al Hattab M and Ghaly A, Microalgae oil extraction pre-treatment methods: critical review and comparative analysis. *J Fundamentals Renew Ener Appl* **5**(4):172 (2015).

21. Daniel M, DWSIM 6.5.3 ed2021.

22. Kooijman HA and Taylor R, ChemSep–another software system for the simulation of separation processes. *CACHE News* **35**:1–9 (1992).

23. Pérez-López P, González-García S, Allewaert C, Verween A, Murray P, Feijoo G *et al.*, Environmental evaluation of eicosapentaenoic acid production by Phaeodactylum tricornutum. *Sci Total Environ* **466**-**467**:991–1002 (2014).

24. Becker EW, *Nutritional Properties of Microalgae: Potentials and Constraints*. CRC Handbook of Microalgal Mass Culture. CRC Press, Boca Raton (1986).

25. Borowitzka MA, High-value products from microalgae—their development and commercialisation. *J Appl Phycol* **25**(3):743–756 (2013).

26. Pirwitz K, Rihko-Struckmann L and Sundmacher K, Comparison of flocculation methods for harvesting Dunaliella. *Bioresour Technol* **196**:145–152 (2015).

27. Wijffels RH, Barbosa MJ and Eppink MHM, Microalgae for the production of bulk chemicals and biofuels. *Biofuel Bioprod Biorefin* **4**(3):287–295 (2010).

Johannes Kopton

Johannes Kopton holds an MSc in systems engineering and engineering cybernetics from Otto-von-Guericke University Magdeburg, Germany. He completed his master's thesis at the Max Planck Institute for Dynamics of Complex Technical Systems in Magdeburg and is now a PhD student in the PhenoRob cluster of excellence at the University of Bonn. His main scientific interests include the optimization of complex systems, decision analysis, and sustainability assessment.

Liisa K Rihko-Struckmann

Liisa K Rihko-Struckmann graduated in chemical engineering at the Helsinki University of Technology (HUT), Finland. She started her career as an industrial catalysis scientist at the Neste Oil Refinery, Finland, in 1989. Parallel to that she pursued scientific research in reaction engineering at HUT, and received a Dr Sci (Technol) in 1997. After a post-doc stay in Helsinki, she moved to the Max Planck Institute for Dynamics of Complex Technical Systems in Magdeburg, Germany. Her main scientific interest fields are the design and modeling of chemical processes, especially for CO_2 conversion and algal biomass valorization. Since 2009, she has lectured at the Otto-von-Guericke University Magdeburg on the topic of biofuels and sustainability assessment, and life cycle analysis (LCA).

Laura König-Mattern

Laura König-Mattern holds a BSc in biosystems engineering and an MSc in cybernetics from Otto-von-Guericke University Magdeburg, Germany. She works on solvent screening and design for biomass processing as a PhD student at the Max Planck Institute for Dynamics of Complex Technical Systems in the group of Prof. Kai Sundmacher. She was a guest PhD in the group of Professor Jeremy Luterbacher at École Polytechnique Fédérale de Lausanne. She is a Christiane Nüsslein-Volhard Fellow, and a BMBF science communication awardee.

Kai Sundmacher

Professor Dr-Ing. Kai Sundmacher studied mechanical engineering and process engineering at the University of Hannover as well as at TU Clausthal (doctorate 1995, habilitation 1998). After a postdoctoral period at the University of Newcastle, UK, in 1999 he was appointed to the Chair of Process Systems Engineering at the University of Magdeburg. Since 2001 he has also been director at the Max Planck Institute for Dynamics of Complex Technical Systems in Magdeburg. His research interests include the modeling, analysis, and synthesis of complex processes in chemical production, biotechnology, and chemical energy conversion.