A Pareto aggregation approach for environmental-economic multi-objective optimization applied on a second-generation bioethanol production model

Reference:
Vasilakou Konstantina, Billen Pieter, Van Passel Steven, Nimmegeers Philippe.- A Pareto aggregation approach for environmental-economic multi-objective optimization applied on a second-generation bioethanol production model
Energy conversion and management - ISSN 1879-2227 - 303(2024), 118184
Full text (Publisher's DOI): https://doi.org/10.1016/J.ENCONMAN.2024.118184
To cite this reference: https://hdl.handle.net/10067/2030460151162165141

A Pareto aggregation approach for environmental-economic multi-objective optimization applied on a second-generation bioethanol production model

Konstantina Vasilakou[a,b,c,*], Pieter Billen[b], Steven Van Passel[a,c,d], Philippe Nimmegeers[a,b,c]

[a]Environmental Economics (EnvEcon), Department of Engineering Management, Faculty of Business and Economics, University of Antwerp, Prinsstraat 13, 2000 Antwerp, Belgium.

[b]Intelligence in Processes, Advanced Catalysts and Solvents (iPRACS), Faculty of Applied Engineering, University of Antwerp, Groenenborgerlaan 171, 2020 Antwerp, Belgium.

[c]Flanders Make@UAntwerp, 2000 Antwerp, Belgium

[d]Nanolab Centre of Excellence, Prinsstraat 13, 2000 Antwerp, Belgium.

E-mail addresses: *konstantina.vasilakou@uantwerpen.be* *(K. Vasilakou),* *pieter.billen@uantwerpen.be* *(P. Billen),* *steven.vanpassel@uantwerpen.be* *(S. Van Passel),* *philippe.nimmegeers@uantwerpen.be* *(P. Nimmegeers).*

ABSTRACT

Multi-objective optimization is an important decision-making tool for energy processes, as multiple targets need to be achieved. These objectives are usually conflicting since a single solution cannot be optimal for all objectives, resulting in a set of Pareto-optimal solutions. Multiple indicators might be available to describe a sustainability objective, such as the environmental impact which is commonly evaluated by performing a life cycle assessment. In this study, Pareto aggregation is proposed as a method which employs a novel multi-objective

*Corresponding Author at: Environmental Economics (EnvEcon), Department of Engineering Management, Faculty of Business and Economics, University of Antwerp, Prinsstraat 13, 2000 Antwerp, Belgium. E-mail address: konstantina.vasilakou@uantwerpen.be

optimization-based approach as an alternative to the classically used aggregation in life cycle assessment. This method identifies conflicting environmental indicators and performs an aggregation among those that require a trade-off. An environmental-economic optimization of a second-generation bioethanol plant is used to illustrate and evaluate the proposed method. Process parameters from a biochemical conversion pathway flowsheet simulation model are chosen as optimization variables. To reduce the computational time, surrogate models, based on artificial neural networks, are used. Out of the eighteen ReCiPe Midpoint environmental indicators, five were identified as conflicting, resulting in an aggregated environmental objective, which was then traded off with the economic objective function, chosen as the levelized cost of ethanol. Comparison with the widely used single-score EcoIndicator99 showed that the Pareto aggregation method can reduce most of the environmental indicators by up to 6.5%. This research provides an insight on non-redundant objective functions, aiming at reducing the dimensionality of multi-objective optimization problems, while taking into consideration decision-makers' preferences.

KEYWORDS

Multi-objective optimization, sustainability optimization, life cycle assessment, bioprocess modelling, biorefinery

NOMENCLATURE

Abbreviations

ANN: Artificial neural network

BP: Best point

CAPEX : Capital expenditure

LCA: Life cycle assessment

MOO: Multi-objective optimization

MSE: Mean squared error

OPEX: Operational expenditure

Symbols

a: annuity factor

C_{FCI}: fixed capital investment (EUR)

$C_{operlab}$: operating labor cost (EUR/h)

C_{opfix}: fixed operating cost (EUR/y)

C_{opvar}: variable operating cost (EUR/y)

$\tilde{J}_i$: predicted value of J_i

${J_i}^{BP}$: objective value of J_i on the Best point

$J_{agg,env}$: aggregated environmental objective

J_{conf}: set of conflicting objectives J_i

J_{env}: set of environmental objectives J_i

J_i: objective function i

$J_i^{BP,EI99}$: objective value of J_i on the Best point of the EcoIndicator99-economic Pareto front

$J_i^{BP,agg}$: objective value of J_i on the Best point of the aggregated environmental-economic Pareto front

J_{obj}: set of objective functions

k: number of objectives for each combination

LCE: levelized cost of ethanol (EUR/L)

N: number of Pareto fronts

n: number of samples

n_J: number of objective functions J

n_{JC}: number of conflicting objectives J_i

n_{Jenv}: number of environmental objectives

P_{EtOH}: annual ethanol production (L/y)

PF_{JC}: number of conflicting Pareto fronts

W_i: weights of each conflicting objective $J_i \in J_{conf}$

x: vector of optimization variables

x_{max}: vector of the maximum values of the optimization variables

x_{min}: vector of the minimum values of the optimization variables

y_i: variable of min-max scaling

y_i': normalized value of y_i variable

U: Utopia point of two objective functions J_i and $J_{j \neq i}$

1. INTRODUCTION

During the design and operation of energy systems, decisions have to be taken considering multiple objectives, namely: maximizing economic performance (e.g., Net Present Value (NPV), profit), while minimizing environmental impact (e.g. global warming potential, carbon footprint, ecosystem quality savings) [1]. These objectives can be conflicting, as improving one can result in worsening the other. In such a situation, there is not one optimal solution, but rather several mathematically equivalent trade-off solutions exist (i.e., Pareto optimal solutions) [2]. Multi-objective optimization (MOO) methods generate such a set of Pareto optimal solutions, called the Pareto front [3], and have been widely used in energy applications [4]. These mathematically equivalent trade-off solutions can then assist in the decision-making process. The systematic generation and efficient presentation of these optimal alternatives to decision-makers plays a key role in computer-aided decision making. Decisions need to be made in an efficient and well-informed manner while the decision-makers' preferences also need to be taken into account [1].

In environmental assessments, such as the widely used life cycle assessment (LCA), many different indicators are exploited to quantify and evaluate environmental impacts [5]. Usually only CO_2 emissions are taken into consideration in multi-objective optimization problems on energy systems [6,7], neglecting the rest of the indicators. On the other hand, multiple environmental indicators can be normalized and weighted, resulting in a single score indicator. The use of a single aggregated indicator facilitates the communication and interpretation of the comparative results from LCA practitioners to decision-makers [8]. For example, endpoint impact categories (i.e., damages to human health, ecosystem quality and resources) are used to have a comprehensive view on environmental impacts [9]. Then, these impact categories are often aggregated into a single Eco-Indicator [10] (e.g., Eco-indicator 99 [9]).

However, the weighted sum procedure has received lot of criticism over the years, due to the danger of incorrectly interpreting the weighted results, as weighting factors are subjective and might not represent decision-makers' preferences [8,11]. As a result, multiple efforts have been made to improve the aggregation approach in environmental assessments. Afrinaldi et al. [8] have proposed a novel method for normalization and aggregation in LCA based on fuzzy logic, which was applied in automotive engines. Similarly, Agarski et al. [12] have used fuzzy logic for impact category weighting in an LCA study on waste treatment processes. The analytical hierarchy process (AHP) has also been applied to estimate weighting factors for an electricity generation case study [11]. Moreover, Sohn et al. [13] have developed a novel weighting method, called Argumentation Corrected Context Weighting-LCA (ArgCW-LCA), which uses multi-criteria decision analysis to aggregate midpoint impacts to a single indicator.

Despite these efforts, the use of aggregated impact indicators as the final environmental objective function in optimization problems, can lead to suboptimality and a loss of information for decision makers [14,15]. Indeed, the attained solution of the problem is influenced by the weighted sum procedure as the conflicting behavior of different impact categories is not considered, leading to suboptimal solutions. A recent study by Zacharopoulos et al. [16] focused on optimizing battery electric vehicles (BEV) charging profiles by minimizing their environmental impact, while also including the identification of conflicting and non-conflicting environmental midpoint impact indicators. The conflicting indicators were determined by first optimizing each impact indicator separately, then calculating the rest of the indicators for these optimized charging profiles and finally measuring the deviation between the objectives. Despite identifying conflicting environmental objectives, the aggregation of these conflicting indicators into a single final environmental objective was not included in the study.

To address this knowledge gap, a novel method named *Pareto aggregation* is proposed in this study which firstly solves systematically a multi-objective optimization problem related to environmental sustainability using different indicators and identifying which indicators are truly conflicting and between which ones a tradeoff needs to be made. In contrast to the current state-of-the-art, by identifying the difference between conflicting and non-conflicting objectives, the objective space is reduced by only retaining the conflicting ones, as the non-conflicting ones will lead to the same optimal solution (and hence a waste of computational power and time). Based on decision-maker preferences and/or the level of conflict between the objectives, an aggregation is performed, resulting in one aggregated environmental objective, which is then traded off against an economic objective function when solving a second bi-objective environmental-economic optimization problem. This method is especially designed to tackle in a systematic way multi-objective optimization formulations which are normally solved via aggregation, although it is generally applicable to all types of MOO problems. Thus, it can serve as a useful tool in decision-making processes, commonly encountered during the design and operation of energy systems, when the technical, exergetic, economic, environmental and/or social performances are often evaluated.

To illustrate this methodology, second-generation bioethanol production is chosen as a realistic and complex energy conversion case study. Biomass has emerged as a renewable energy resource with a high potential in the worldwide efforts for a greener energy transition [17]. Out of its various applications, the production of biofuels can assist in achieving future climate targets and meeting energy demand [18]. In particular, lignocellulosic biomass is an abundant carbon source, rich in energy components that can be converted to biofuels, commonly known as second-generation biofuels [19]. Due to its recalcitrant structure, the bioconversion of lignocellulose requires multiple complex processes [20]. The optimization of such energy conversion systems has been an area of interest for the past years, with numerous studies focusing on the optimal process design of biofuels production with respect to economic and environmental criteria [21]. Dynamic process models are developed for the upstream processes, those being dilute acid pretreatment, enzymatic hydrolysis and fermentation, by using already developed kinetic models. Due to their complexity and long computational time, surrogate models are developed and used instead [22], calculating the final optimization objectives. In addition to this case study, the Pareto aggregation method can easily be applied to different types of MOOs and systems. Its applicability is explained further in subsection 3.4.1, along with specific directions and requirements.

2. MATERIAL AND METHODS

The proposed methodology consists of three steps, indicated by A, B and C in Figure 1. First, rigorous process models are developed to calculate mass and energy balances which are the basis for the environmental and economic objective functions (Figure 1 (A)). For the lignocellulosic bioethanol production these models are developed in ASPEN Plus, by using kinetic models available in literature. Through an interface connection between ASPEN Plus and MATLAB, the final economic and environmental objectives are calculated. Due to the computationally expensive interface between ASPEN Plus and MATLAB and time required to solve multi-objective optimization problems, artificial neural networks (ANN) are developed as surrogate models for each objective function (Figure 1 (B)). As such the computational cost to evaluate objective functions is reduced and the optimizations can be conducted efficiently. The development of surrogate models is also highly relevant when more complicated models are used that require more computational time. Finally, the Pareto aggregation method is applied in order to calculate the final aggregated environmental objective, which is then traded off against the economic objective (Figure 1 (C)) through a multi-objective optimization. A detailed description of each methodology step is given in the next subsections.

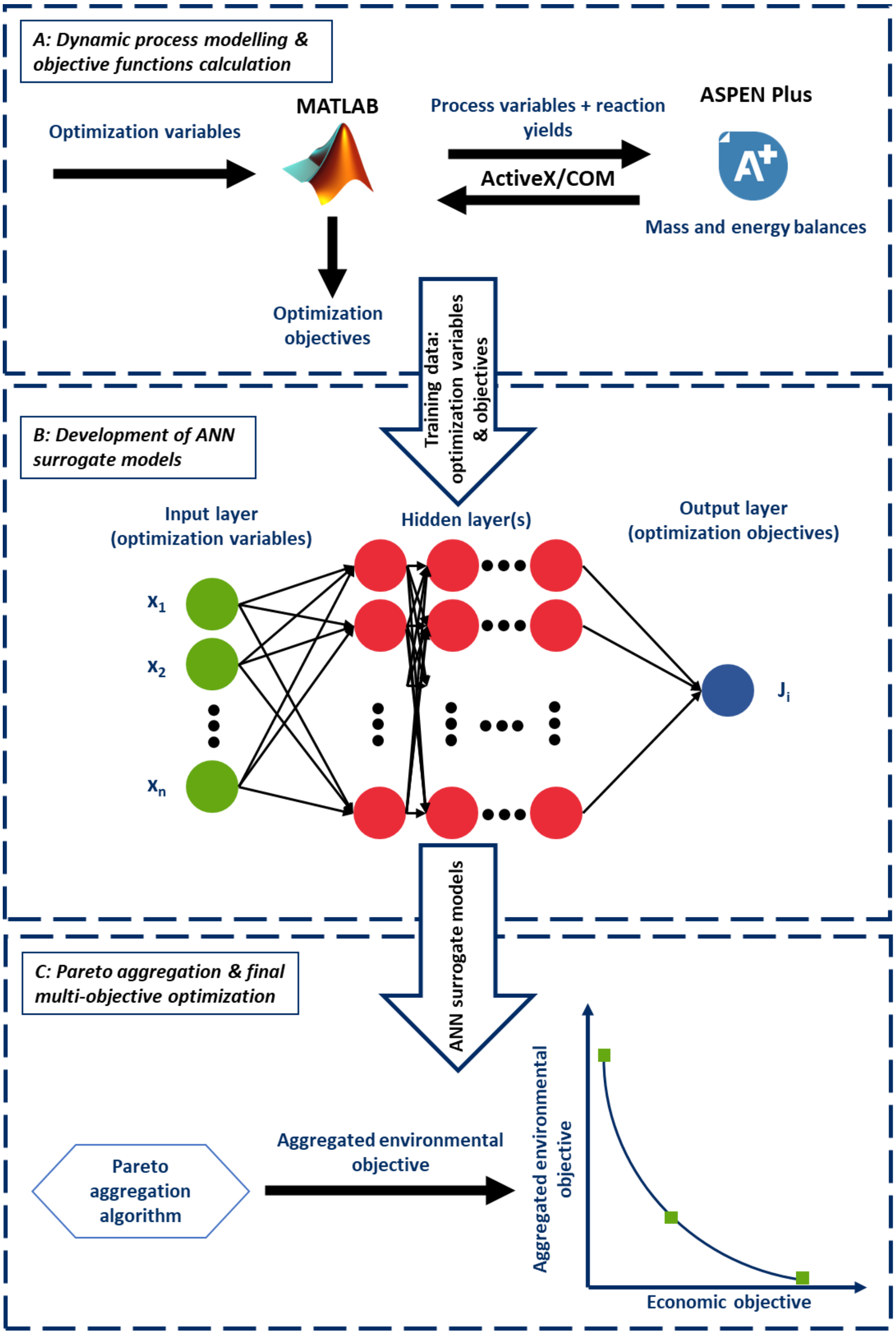

Figure 1. Schematic overview of the methodology steps applied: (A) Dynamic process modelling and calculation of objective functions, (B) Development of ANN surrogate models, (C) Pareto aggregation and multi-objective optimization.

2.1 Rigorous process model development & simulations

Bioethanol production from corn stover is simulated in ASPEN Plus® v.12.1 [23], based on the model of the National Renewable Energy Laboratory (NREL) [24]. Three main upstream processes are included: dilute acid pretreatment, enzymatic hydrolysis and co-fermentation. Dilute acid pretreatment was modelled according to Shi et al. [25] for hemicellulose degradation, while the solubilization of lignin according to Lavarack et al. [26]. Following the model of Humbird & Aden [27], a two-stage pretreatment process was assumed with 70% conversion of oligomers to monomers. The kinetic model of Kadam et al. [28] was used for the enzymatic hydrolysis (saccharification), assuming three main reactions with competitive inhibition. Finally, co-fermentation by recombinant Zymomonas mobilis was assumed for the final process, applying the kinetic model of Leksawasdi et al. [29].

Thus, all reaction yields are estimated based on the kinetic models used for each process (see Supporting Material). The developed Ordinary Differential Equation (ODE) and Differential Algebraic Equation (DAE) systems are solved in MATLAB R2022b with the built-in ode15s function, transferring the calculated reaction yields to ASPEN, through an ActiveX interface connection.

2.2 Economic indicator calculation

The levelized production cost of ethanol (LCE, EUR/L) is chosen as the main economic indicator, calculated using equation (1) [30]:

$$LCE = \frac{(CAPEX \cdot \alpha + OPEX)}{P_{EtOH}} \tag{1}$$

Where CAPEX is the capital expenditure (EUR), OPEX the operational expenditure (EUR/y), α the annuity factor (y^{-1}) and P_{EtOH} the annual ethanol production (L/y). The annuity factor is calculated assuming 20 years of project lifetime and 15% discount rate [30], for a production plant starting its operation in 2022.

The CAPEX is calculated by summing the fixed capital investment C_{FCI} (EUR), the working capital (5% of the C_{FCI}) and the land cost (2% of the C_{FCI}) [24]. The C_{FCI} consists of the total direct and indirect costs. The OPEX is calculated as the sum of the variable operating (C_{opvar}) and the fixed operating (C_{opfix}) costs. Details on the input values and calculations are given in the Supporting Material.

2.3 Environmental indicators calculation

The environmental performance is evaluated through a life cycle assessment (LCA). The system covers the cultivation and supply of biomass, the supply of raw materials and energy as well as the production of the final product. A cradle-to-gate approach is chosen, as the case study is limited to the upstream processes of lignocellulosic ethanol (EtOH) production. The functional unit is taken as 1 L of ethanol production. For the biomass cultivation, an economic allocation is applied (11.3% for corn stover).

The EcoInvent database [31] and data on Belgian corn agriculture are used to develop the life cycle inventory (see Supporting Material). The input and output flows are taken from the ASPEN model and are expressed per L EtOH (i.e. the functional unit). The life cycle impact assessment is performed in SimaPro® v.9.4.02, using the ReCiPe 2016 v1.1 Midpoint method [32] and calculating eighteen environmental impact indicators.

For the validation of the suggested methodology, the EcoIndicator99 is also calculated, using the ReCiPe 2016 v1.1 Endpoint method and its normalization factors, while weights are taken from the methodology report on the EcoIndicator 99 [9]. This single score was chosen as it is well recognized and widely used in relevant studies.

2.4 Multi-objective optimization problem formulation

Multi-objective optimization problems of the following form are studied in this work:

$$\min_{x}\left\{J_1(x,p),\dots,J_{n_J}(x,p)\right\} \tag{2}$$

subject to

$$c(x,p) \leq 0$$

$$x_{min} \leq x \leq x_{max}$$

With $J_{obj} = \{J_1(x,p),\dots,J_{n_J}(x,p)\}$ the n_J objective functions, $c(x,p)$ the constraint functions, x the vector with optimization variables, p the model parameter vector, x_{min} the vector containing the minimum values of the optimization variables and x_{max} the vector with the maximum values of the optimization variables.

The solution of such a multi-objective optimization problem is a Pareto front of tradeoff solutions. The Non-dominated Sorting Genetic Algorithm II (NSGA-II), an elitist evolutionary algorithm, is used in this work. Problems are solved in MATLAB R2022b using the *gamultiobj* function, taking function tolerance at 10^{-4} and constraint tolerance at 10^{-12}. All calculations were performed using the Intel® Xeon® Gold 6334 CPU @ 3.60GHz – 3.59GHz (4 processors) and 16.0 GB RAM.

2.5 Surrogate modelling

Surrogate models are used for the calculation of the objective functions, as the coupling of ASPEN and MATLAB is computationally expensive. Artificial Neural Networks (ANN) are developed in this study for each objective function [33]. Thus, a set of 20 models is created, as shown in Table 1.

Table 1. Objective functions for the multi-objective optimization of second-generation bioethanol production. J_1 is the economic indicator, J_2-J_{19} are the environmental indicators and J_{20} is the EcoIndicator used for comparison.

Optimization objectives	Objective functions	Units
LCE	J_1	EUR/L
Global warming	J_2	kg CO_2 eq/L
Stratospheric ozone depletion	J_3	kg CFC11 eq/L
Ionizing radiation	J_4	kBq Co-60 eq/L
Ozone formation, Human health	J_5	kg NO_x eq/L
Fine particulate matter formation	J_6	kg PM2.5 eq/L
Ozone formation, Terrestrial ecosystems	J_7	kg NO_x eq/L
Terrestrial acidification	J_8	kg SO_2 eq/L
Freshwater eutrophication	J_9	kg P eq/L
Marine eutrophication	J_{10}	kg N eq/L
Terrestrial ecotoxicity	J_{11}	kg 1,4-DCB/L
Freshwater ecotoxicity	J_{12}	kg 1,4-DCB/L
Marine ecotoxicity	J_{13}	kg 1,4-DCB/L
Human carcinogenic toxicity	J_{14}	kg 1,4-DCB/L
Human non-carcinogenic toxicity	J_{15}	kg 1,4-DCB/L
Land use	J_{16}	m^2a crop eq/L
Mineral resource scarcity	J_{17}	kg Cu eq/L
Fossil resource scarcity	J_{18}	kg oil eq/L
Water consumption	J_{19}	m^3/L
EcoIndicator 99	J_{20}	Points/L

2.5.1 Sampling and generation of training data

The training data required for the surrogate models are obtained by sampling the design space with the Best Candidate algorithm [34]. The chosen algorithm performs well and satisfies major requirements for surrogate modelling applications [35].

Feed rate, acid loading, pretreatment temperature, pretreatment residence time, saccharification residence time and fermentation residence time are the optimization variables chosen. The six-variable design space (6D) is sampled at once for 5000 samples, according to the permitted lower and upper bounds of Table 2. These limits are selected based on the constraints of the kinetic models used for the process modelling, while ensuring that the simulation model runs without errors (simulation status was checked after each run). The generated samples are then used to run the simulation model and calculate the economic and environmental indicators. Since both inputs and outputs have different scales that affect the sensitivity and convergence of the developed surrogate models, they are both normalized from 0 to 1, using the min-max scaling equation (3):

$$y_i' = \frac{y_i - \min(y_i)}{\max(y_i) - \min(y_i)} \tag{3}$$

Where y_i' is the normalized value of y_i variable.

Table 2. Upper and lower limits of optimization variables.

Optimization variables (x)	Lower bound (x_{min})	Upper bound (x_{max})
x_1: Acid loading (mg/g dry biomass)	1	2
x_2: Pretreatment temperature (°C)	155	185
x_3: Pretreatment residence time (min)	1	20
x_4: Feed (dry t/d)	80	2000
x_5: Saccharification residence time (min)	10	120
x_6: Fermentation residence time (min)	10	50

2.5.2 ANN development

First, the architecture of the ANN is studied and optimized for each objective. In this study, only shallow and two-hidden layers ANNs are considered due to their simplicity. The number and size of layers are optimized, taking a maximum number of neurons per layer as double the amount of inputs based on rules-of-thumb [36], that being 12. The mean squared error (MSE) is used as the performance criterion to select the optimal ANN architecture:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(J_i - \tilde{J}_i)^2 \tag{4}$$

Where n is the number of samples, J_i is the real value and $\tilde{J}_i$ is the predicted value.

The ANN models are developed in MATLAB R2022b, using the *fitnet* function of the Deep Learning Toolbox. The training data are partitioned into training, validation and testing sets at a 70%, 20% and 10% ratio respectively. The Levenberg-Marquardt training method is chosen, while the rest of the training parameters are kept the same as the default options.

2.6 Pareto aggregation algorithm

The proposed Pareto aggregation algorithm is presented in Figure 2. First, multiple bi-objective Pareto fronts are computed for each unique combination of two environmental objectives. The number of Pareto fronts N is calculated by equation (5):

$$N(n_{Jenv}, k) = \frac{n_{Jenv}!}{k!\,(n_{Jenv} - k)!} \tag{5}$$

Where n_{Jenv} is the number of environmental objectives and k is the number of objectives for each combination. In this work, the environmental objectives are $n_{Jenv} = 18$: $J_{env} = \{J_i \mid i = 2, \dots, 19\}$. For 18 environmental indicators taken two at a time, 153 Pareto fronts are required.

Then, the most conflicting objectives are identified based on specific criteria, and weights are generated as described in the following subsections. This way the optimization is focused on finding solutions that balance the conflicting objectives first, leading to more robust results. Finally, the aggregated environmental objective is formed and traded-off against the economic objective J_1.

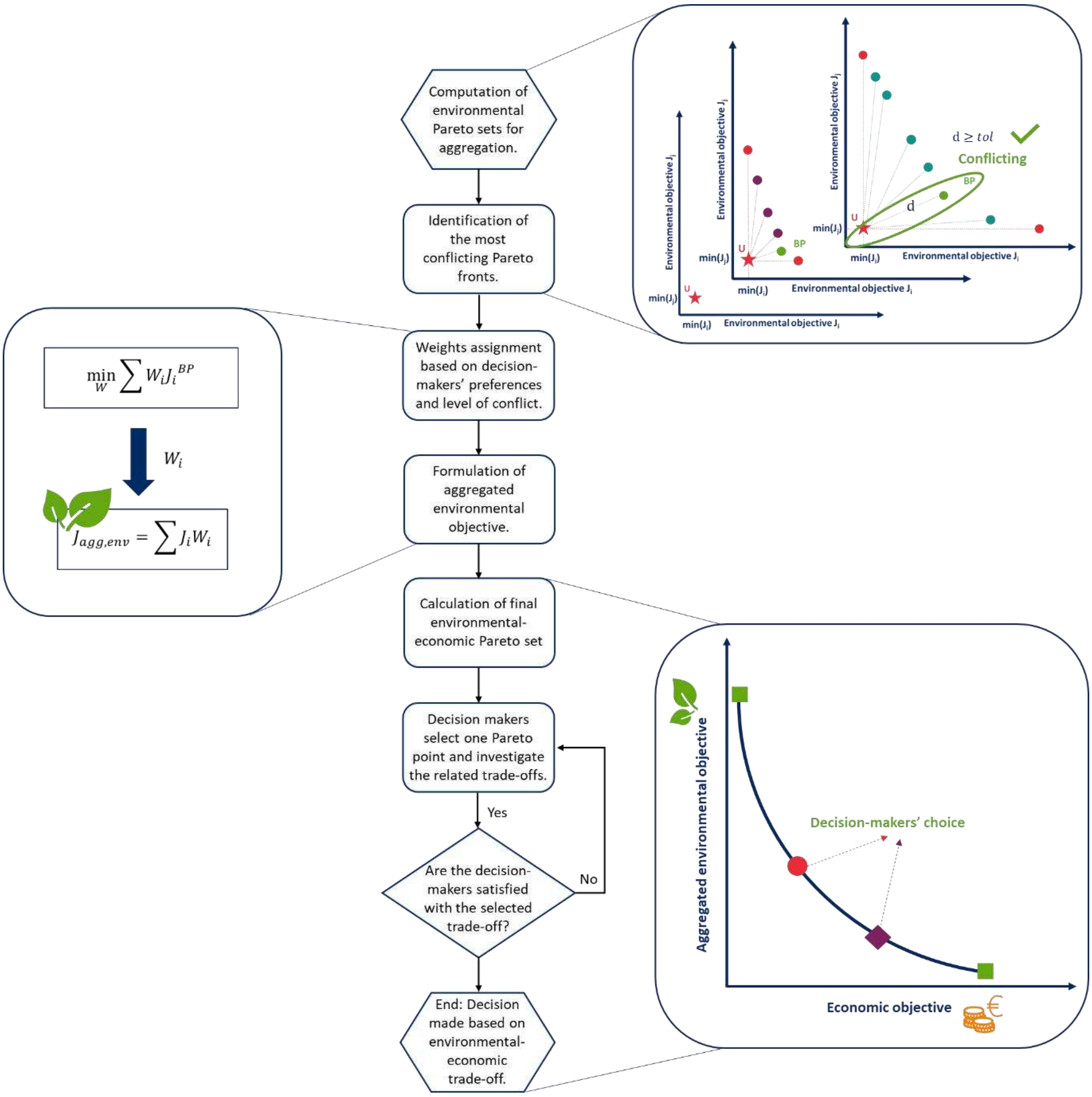

Figure 2. Schematic representation of the Pareto aggregation algorithm (BP: Best point, U: Utopia point, d: Euclidean distance between BP and U, tol: tolerance value, W: Weights used for aggregation).

2.6.1 Identification of conflicting objectives

The most conflicting objectives $J_i \in J_{env}$ are identified using the following criteria:

1. The Pareto front should consist of more than one point.

2. The minimum Euclidian distance of the Pareto points from the Utopia point should be greater than a tolerance value:

$$\min\left\{\sqrt{|U-J_i|^2+|U-J_{j\neq i}|^2}\right\} \geq tol, \qquad J_i, J_{j\neq i} \in J_{env} \tag{6}$$

Where $U = (\min(J_i), \min(J_{j\neq i}))$ is the Utopia point and tol is the tolerance value. The Utopia point consists of the individual minima of each objective function. The tol parameter can be specified by the decision-maker, reflecting what is defined as conflicting. Indeed, its value can be varied to make the criterion more or less strict. In this study, a tolerance of 10^{-3} is deemed as suitable, as the objective values are normalized from 0 to 1. The point of each Pareto front that satisfies the left part of equation (6) is hereby mentioned as "Best point" (BP). A simplified example is given in Figure 3.

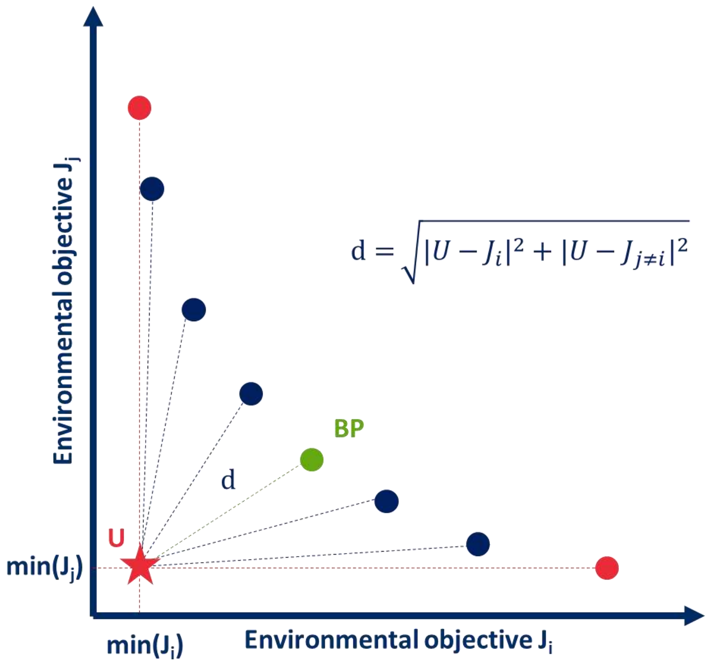

Figure 3. Environmental Pareto front for aggregation between objectives J_i and J_j. U is the Utopia point, BP the Best point and d the distance of each Pareto point from U.

By satisfying both of these criteria, the most conflicting objectives (J_{conf}) are identified, with $J_{conf} \subseteq J_{env}$. The number of conflicting objectives is n_{JC} corresponding to PF_{JC} Pareto fronts.

2.6.2 Weights generation

The final aggregation of the most conflicting environmental objectives can be done through a weighted sum method:

$$J_{agg,env} = \sum_{i \in [2,19] | J_i \in J_{conf}} J_i W_i \tag{7}$$

Where W_i is the weight of the conflicting objective $J_i \in J_{conf}$, with $\sum_{i \in [2,19] | J_i \in J_{conf}} W_i = 1$.

For the generation of the weights, a novel approach is suggested that accounts for decision-makers' preferences (if available) and/or the level of conflict between the objectives. Higher weights can be assigned to objectives that are more important to decision-makers and are highly conflicting:

$$\min_{W} \sum_{i \in [2,19] | J_i \in J_{conf}} W_i J_i^{BP} \tag{8}$$

$$\text{s.t.}$$

$$\sum_{i \in [2,19] | J_i \in J_{conf}} W_i = 1$$

$$A \cdot W \leq b$$

$$c(W) \leq 0$$

$$0.01 \leq W \leq 1$$

Where W is a vector containing the weights, J_i^{BP} are the objective values on the Best Point (BP) of $J_i \in J_{conf}$, A is a matrix, b a vector and $c(W)$ a nonlinear inequality function that returns a vector. The values of A and b can be chosen in order to reflect the decision-makers' preferences and/or the level of conflict between the objectives. The nonlinear inequalities are taken as $c(W) = \left({J_i}^{BP} - W_i {J_i}^{BP}\right)^2 - 10^{-8} \leq 0$, ensuring that the weight generation is not significantly influenced by the absolute values of the objectives (e.g. extremely high weights assigned to low objective values). For the same reason, the lower bound of the weights is taken as 0.01 in order to guarantee that high objective values ${J_i}^{BP}$ are not given extremely low weights, close to 0 (e.g. 10^{-6}-10^{-4}). The *fmincon* MATLAB function is applied, using the default options of the Interior Point Method.

3. RESULTS AND DISCUSSION

3.1 Surrogate modelling results

The 6D design space was sampled 5000 times using the Best Candidate algorithm [34], requiring 470 seconds of running time. Detailed results on its performance are available in the Supplementary Material. ANNs were then used to create 20 surrogate models in total between the optimization variables and the optimization objectives; one for each out of the 19 ReCiPe environmental indicators and one for the EcoIndicator99. The optimal architecture, identified through the calculation of MSE, was found for each surrogate model, as shown in Table 3. Mean squared errors in the magnitude of 10^{-5}-10^{-6} were achieved for all models for a two-layered ANN. The coefficients of determination (R^2) for the training, validation, testing total data are also shown in Table 3. These are higher than 99% for all models and data, indicating good predictions and that surrogate models are able to capture the relationships between the objective functions and optimization variables.

Table 3. Performance and architecture of the optimal ANN for each surrogate model/objective function. R^2 train, R^2 val, R^2 test and R^2 total are the coefficients of determination for the training, validation, testing and total data respectively.

Surrogate model/Objective function	MSE	R^2 train	R^2 val	R^2 test	R^2 total	No of neurons in 1st layer	No of neurons in 2nd layer
J_1	8.36e-06	0.9992	0.9995	0.9997	0.9993	10	8
J_2	1.70e-05	0.9999	0.9976	0.9997	0.9995	11	9
J_3	9.65e-06	0.9993	0.9998	0.9997	0.9995	12	10
J_4	2.58e-06	0.9998	0.9986	0.9997	0.9995	12	6
J_5	1.40e-05	0.9999	0.9998	0.9956	0.9995	11	7
J_6	1.47e-05	0.9992	0.9997	0.9996	0.9995	10	6
J_7	1.38e-05	0.9992	0.9998	0.9998	0.9995	11	9
J_8	1.64e-05	0.9992	0.9994	0.9997	0.9995	11	7
J_9	1.29e-05	0.9993	0.9997	0.9998	0.9995	12	7
J_{10}	1.23e-05	0.9993	0.9998	0.9997	0.9995	11	6
J_{11}	1.29e-05	0.9993	0.9998	0.9998	0.9995	12	8
J_{12}	7.82e-06	0.9994	0.9998	0.9998	0.9995	12	10
J_{13}	1.07e-05	0.9998	0.9997	0.9955	0.9995	11	8
J_{14}	9.56e-06	0.9993	0.9998	0.9997	0.9995	12	5
J_{15}	9.99e-06	0.9993	0.9997	0.9997	0.9995	11	9
J_{16}	1.12e-05	0.9999	0.9979	0.9999	0.9995	12	7
J_{17}	7.74e-06	0.9999	0.9981	0.9998	0.9995	12	9
J_{18}	1.45e-05	0.9999	0.9998	0.9962	0.9995	12	10
J_{19}	1.64e-05	0.9999	0.9974	0.9984	0.9995	12	9
J_{20}	1.25e-05	0.9999	0.9977	0.9998	0.9994	12	9

3.2 Pareto aggregation results

In total, 153 Pareto fronts were calculated for 18 environmental indicators. Out of these, four ($PF_{JC} = 4$) were found to be conflicting based on the criteria specified in the Pareto aggregation algorithm. In Figure 4, the minimum distance calculated between the Pareto points and the Utopia point (U) is depicted, that being equal or higher to 10^{-3}. Therefore, five environmental indicators are identified as the most conflicting, those being the stratospheric ozone depletion (J_3), ionizing radiation (J_4), terrestrial ecotoxicity (J_{11}), land use (J_{16}) and water consumption (J_{19}).

It is evident from Figure 4 that land use (J_{16}) is highly conflicting with stratospheric ozone depletion (J_3) and terrestrial ecotoxicity (J_{11}), while water consumption (J_{19}) has the least conflicting relation with stratospheric ozone depletion (J_3). Based on these observations and given the lack of decision-makers' preference in the current case study, the linear inequalities of equation (8) are chosen, assigning higher weights to the most conflicting objectives, those being J_{16}, followed by J_3 and J_{11}.

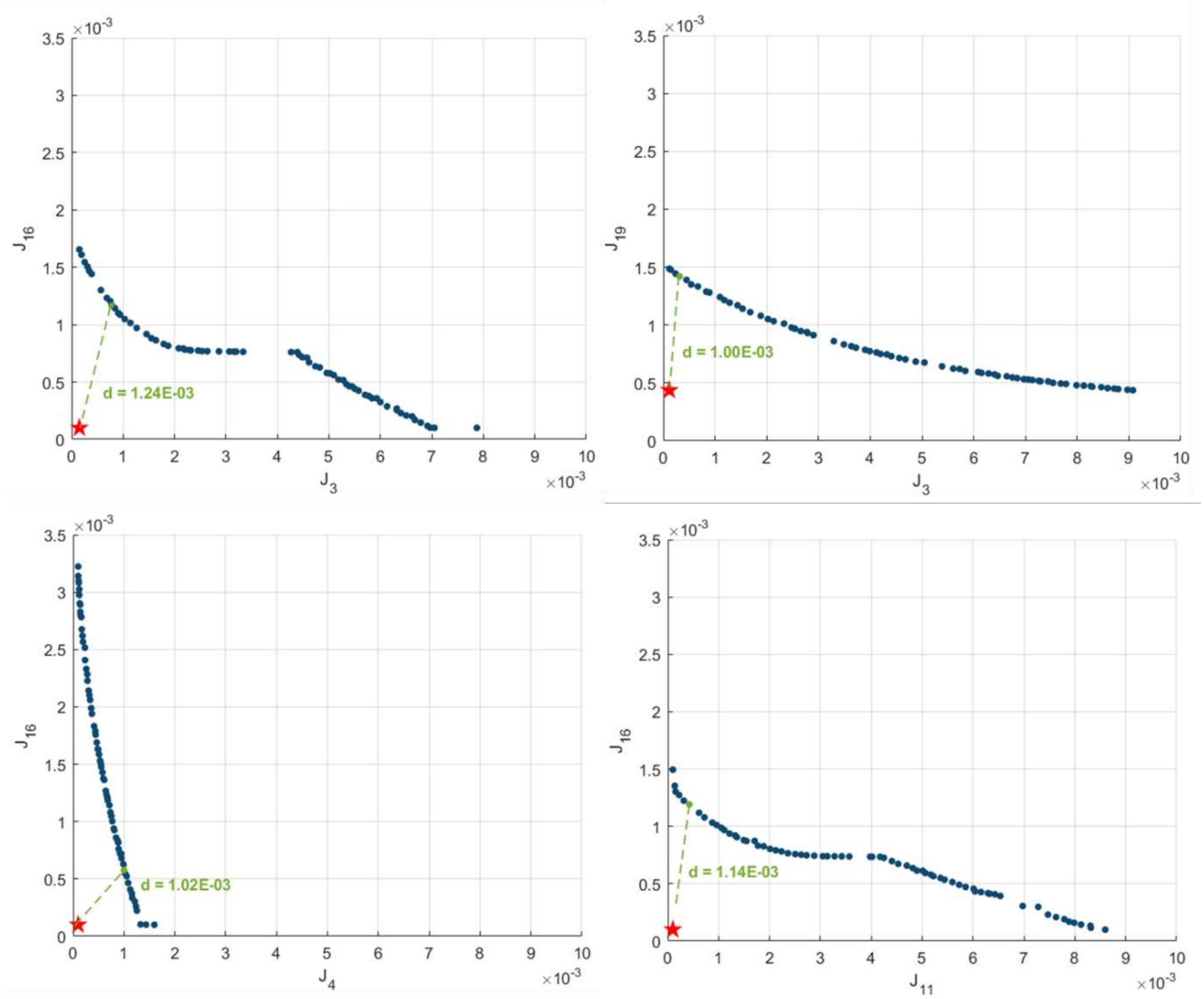

Figure 4. Conflicting objectives identified through the Pareto aggregation algorithm: (A) Stratospheric ozone depletion (J_3) vs Land use (J_{16}), (B) Stratospheric ozone depletion (J_3) vs Water consumption (J_{19}), (C) Ionizing radiation (J_4) vs Land use (J_{16}) and (D) Terrestrial ecotoxicity (J_{11}) vs Land use (J_{16}). The distance (d) between the Utopia Point (★) and the Best Point (●) is also shown. Both objective values are normalized within [0,1].

The identified conflicting behaviour can be explained by the fact that corn stover has the highest emission factor for land use, which on the other hand has an insignificant emission factor for stratospheric ozone depletion, ionizing radiation and terrestrial ecotoxicity, which are mostly affected by the diammonium phosphate and electricity supply. Indeed, ozone depletion and ionizing radiation get both minimized at a very low fermentation residence time, almost three times less than the one for land use, due to the high impact of diammonium phosphate consumption. Similarly, water consumption indicator is mostly affected by the wastewater output, resulting in around 30% less acid loading required for the minimization of this indicator compared to the rest.

The final weights for each objective are presented in Table 4. Land use has the highest contribution to the final environmental objective, followed by stratospheric ozone depletion, terrestrial ecotoxicity, ionizing radiation and water consumption. The performance and reliability of the proposed weight generation algorithm has been verified through a robustness check (details in Supporting Material).

Table 4. Weights assigned to the most conflicting environmental objectives

Environmental objective	Weight
Stratospheric ozone depletion (J_3)	0.3159
Ionizing radiation (J_4)	0.0107
Terrestrial ecotoxicity (J_{11})	0.2546
Land use (J_{16})	0.3188
Water consumption (J_{19})	0.1000
Sum	1.0000

3.3 Environmental-economic multi-objective optimization results

The aggregated environmental objective $J_{agg,env}$ was then traded-off against the economic objective J_1. The final Pareto front is presented in Figure 5(A), obtained for 1200 maximum number of iterations and population size taken as 200. A conflicting relationship between the chosen economic objective and the aggregated environmental objective is identified. Detailed results of the optimization variables can be found in the Supporting Material.

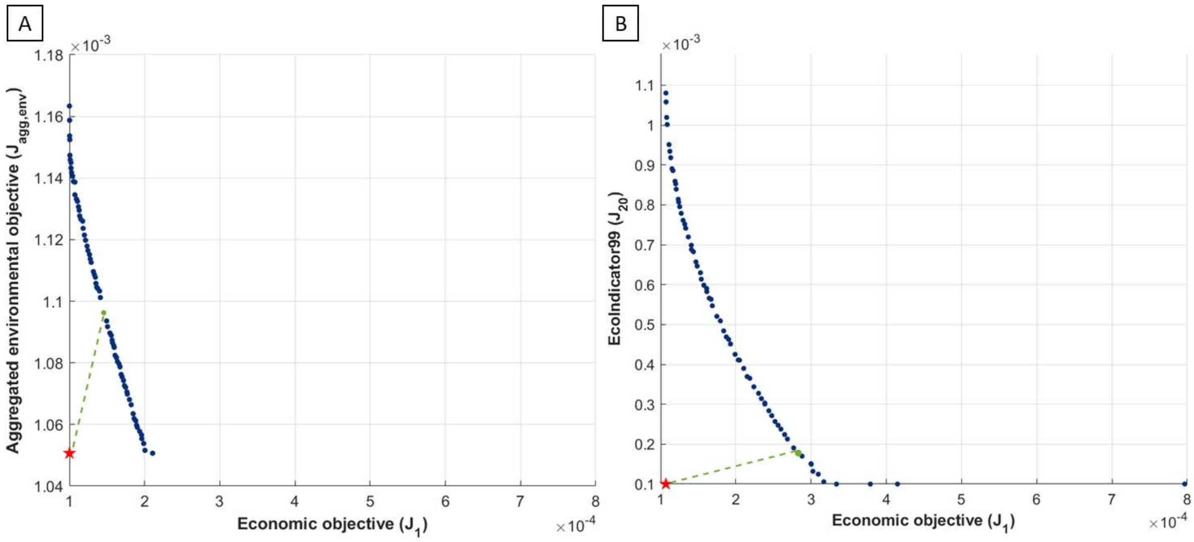

Figure 5. Final environmental-economic Pareto front of the (A) Aggregated environmental objective ($J_{agg,env}$) and (B) EcoIndicator99 (J_{20}). The Utopia Point (★) and the Best Point (●) are also depicted. All objectives are normalized within [0,1].

As far as the optimization variables are concerned, a high biomass feed is required for all of the Pareto optimal solutions, around 1800 dry t/d biomass on average, as the scale of the biorefinery has a significant effect in the economic performance (economies of scale) [37]. A high acid loading (~2.0 mg/g dry biomass) accompanied by low temperature (~159 ºC) and high residence time (~20 min) is required for the acid pretreatment process. The low pretreatment temperature influences both the economic and environmental performance, as less inhibitors are produced while limiting the energy consumption. A high saccharification time (~120 min) is also required in order to achieve a high sugar conversion, while the fermentation time is significantly lower, varying from 35 to 39 min. This can be explained by the fact that after 30 min most of the sugars are already converted to ethanol [29].

In order to validate the aggregation approach proposed in this study, a comparison with a commonly used environmental single-score indicator, EcoIndicator99, has been conducted.

The Pareto front obtained for trading the economic objective against the EcoIndicator99 is presented in Figure 5 (B).

The optimization objective values (non-normalized) on the Best Point of each Pareto front were calculated, allowing for a more comprehensible comparison. The relative difference between the two Pareto fronts for all optimization objectives is presented in Figure 6. It is now evident that almost all of the environmental objectives are lower by applying the Pareto aggregation algorithm compared to the EcoIndicator99, except for the Land Use (J_{16}) which is however insignificantly higher, less than 1%. Notably, Fossil resource scarcity (J_{18}) can be reduced by 6.5%, Freshwater ecotoxicity (J_{12}) by 5.1% and Marine ecotoxicity (J_{13}) by 4.8% by applying the Pareto aggregation algorithm. For both MOOs, the same value on the economic objective (LCE) was obtained, that being 1.25 EUR/L.

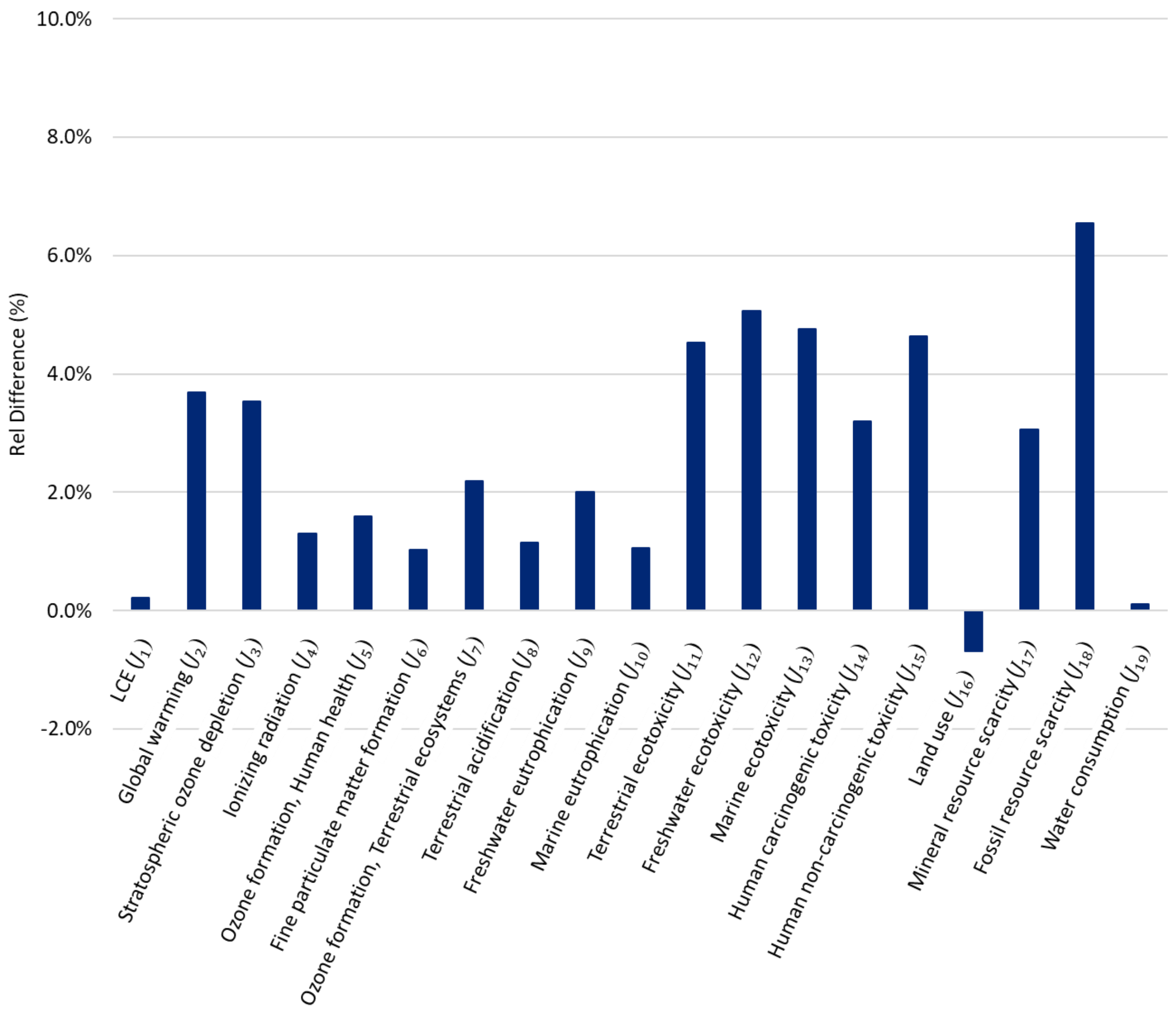

Figure 6. Relative difference between optimization objectives on the Best Point (BP) of the Economic-Aggregated environmental objective $\left(J_i^{BP,agg}\right)$ and the Economic-EcoIndicator99 $\left(J_i^{BP,EI99}\right)$ Pareto fronts. Relative difference calculated as: $\frac{J_i^{BP,EI99} - J_i^{BP,agg}}{J_i^{BP,agg}}$.

These final environmental objectives on the Best point can be attained by decreasing the acid pretreatment temperature by 7°C, while increasing the feed by 47 dry t/d and the fermentation time by 3 min, compared to the EcoIndicator99 results. The rest of the optimization variables are almost the same. Detailed results on the optimization variables and objectives for both Pareto fronts can be found in the Supplementary Material.

3.4 Discussion

The Pareto aggregation method suggested in this study is especially designed to tackle in a systematic way MOO problems which are normally solved via aggregation. This is done by identifying truly conflicting optimization objectives and performing an aggregation of those while taking into consideration their level of conflict and/or decision-makers' preferences, if available.

This method allows the practitioner to account for multiple indicators that are available to describe a single performance (e.g., economic, environmental, social). Nevertheless, the identification of conflicting indicators before the final multi-objective optimization can significantly reduce the dimensionality of the optimization problem. That was indeed illustrated by the applied case study on a lignocellulosic biorefinery, as only five out of the eighteen in total environmental indicators were found to be conflicting and an aggregation between those was required.

Moreover, the criteria used for the identification of the most conflicting objectives are subjected to the practitioners' priorities, and can be adjusted to make the algorithm more or less rigorous. Thus, the tolerance value required by the algorithm (equation (6)) should be carefully chosen based on the case study and desired outcome. Similarly, the weight generation approach is highly dependent on particular parameters specified by the practitioner. It is thus evident that another advantage of the proposed method is that it can easily be adjusted to take into consideration and reflect both practitioners' and decision-makers' preferences.

3.4.1 Method applicability

The developed method was illustrated on an environmental-economic multi-objective optimization problem, as the aggregation issue involved in environmental assessments has been widely discussed in literature [38]. However, its usage can be further expanded and applied to different types of multi-objective optimization problems that require an aggregation of multiple indicators. Based on the obtained results, decisions can be made on the operating conditions of a lignocellulosic bioethanol plant, to minimize the final production cost, while, at the same time, supporting a more environmentally sustainable performance. For the chosen case study, a large plant capacity in addition to low pretreatment time and high saccharification time can achieve a large ethanol production, and thus a low production cost (economies of scale), accompanied by a low environmental impact (less inhibitors production). It can thus serve as a quantitative tool in decision-making processes, exploring the relationship between different objectives and guiding decision-makers to select optimal operating parameters based on their own priorities. Such decision-making processes are critical for energy technologies, covering both energy generation and management, as technical, economic, environmental and social objectives need to be satisfied [4].

The suggested Pareto aggregation method can easily be applied to different case studies. The only requirement is the development of a dynamic model that describes the relationship between the optimization variables and the optimization objectives. For process engineering, this relationship is usually expressed through detailed mass & energy balance calculations. For complex processes, process simulation software, such as ASPEN Plus used in this study, are commonly used to facilitate these calculations. Then, the Pareto aggregation method, i.e. the third step (C) in Figure 1, can be easily applied. In case of computationally expensive models, parallel computing and/or surrogate models (e.g., ANNs used in this study) could help reduce running times.

3.4.2 Limitations & Future work

It is worth mentioning that any uncertainties related to the models applied were not taken into consideration and left out of the scope of this study. However, these uncertainties do not have a direct effect on the suggested method itself, but rather on the final case study-specific results. In future work, the proposed Pareto aggregation algorithm could also be embedded in an interactive multi-objective optimization framework, and the inherent uncertainties of the models should also be taken into account, as these may affect the process [39]. The inclusion of uncertainties could be done by using uncertainty propagation techniques such as linearization, sigma points (unscented transformation) and polynomial chaos expansion to propagate the uncertainty on the model parameters towards the objective functions and constraint functions, as explained by e.g., Nimmegeers et al. [40], and Mores et al. [41]. Nevertheless, the application of the suggested method to different case studies, as explained in section 3.4.1, could help to further improve and operationalize it.

4. CONCLUSIONS

A novel method, named Pareto aggregation, is suggested for generating an environmental objective function, by identifying the most conflicting environmental indicators. This method was applied to a bioethanol production plant in an economic-environmental multi-objective optimization. Surrogate models were developed based on simulation and kinetic models. Five environmental indicators, namely stratospheric ozone depletion, ionizing radiation, terrestrial ecotoxicity, land use and water consumption, were identified as conflicting. Based on the level of conflict, an aggregated environmental objective was developed and traded-off against an economic objective, that being the levelized cost of ethanol. The final Pareto optimal solutions obtained indicate the best performance possibilities for the investigated biorefinery. The method was compared against the single-score EcoIndicator99, demonstrating a better performance as a decrease ranging from 1.0 to 6.5% was observed for almost all indicators calculated through the Pareto aggregation method. This approach can significantly reduce the multi-dimensionality of optimization problems and can be easily applied to other energy systems, serving as a useful tool in decision-making processes when multiple objectives need to be satisfied..

ACKNOWLEDGMENTS

This study was carried out within the framework of the ADV_BIO project financed by the FOD Economie - Energietransitiefonds/ SPF Économie - Fonds de Transition Energétique, call 2019 - 2020 subsidies. Philippe Nimmegeers holds a FWO senior postdoctoral fellowship (grant number: 1215523N) granted by FWO Vlaanderen/Research Foundation Flanders.

APPENDIX A. SUPPLEMENTARY DATA

Supplementary data to this article can be found online: <Supplementary Material>

REFERENCES

[1] Rangaiah G, Sharma S, Sreepathi BK. Multi-objective optimization for the design and operation of energy efficient chemical processes and power generation. Curr Opin Chem Eng 2015;10:49–62. https://doi.org/10.1016/j.coche.2015.08.006.

[2] Schweidtmann AM, Clayton AD, Holmes N, Bradford E, Bourne RA, Lapkin AA. Machine learning meets continuous flow chemistry: Automated optimization towards the Pareto front of multiple objectives. Chem Eng J 2018;352:277–82. https://doi.org/10.1016/j.cej.2018.07.031.

[3] Cheraghi R, Hossein Jahangir M. Multi-objective optimization of a hybrid renewable energy system supplying a residential building using NSGA-II and

MOPSO algorithms. Energy Convers Manag 2023;294:117515. https://doi.org/10.1016/j.enconman.2023.117515.

[4] Al Moussawi H, Fardoun F, Louahlia-Gualous H. Review of tri-generation technologies: Design evaluation, optimization, decision-making, and selection approach. Energy Convers Manag 2016;120:157–96. https://doi.org/10.1016/j.enconman.2016.04.085.

[5] Čuček L, Klemeš JJ, Kravanja Z. A Review of Footprint analysis tools for monitoring impacts on sustainability. J Clean Prod 2012;34:9–20. https://doi.org/https://doi.org/10.1016/j.jclepro.2012.02.036.

[6] Cheraghi R, Hossein Jahangir M. Multi-objective optimization of a hybrid renewable energy system supplying a residential building using NSGA-II and MOPSO algorithms. Energy Convers Manag 2023;294:117515. https://doi.org/10.1016/j.enconman.2023.117515.

[7] Das BK, Hassan R, Tushar MSHK, Zaman F, Hasan M, Das P. Techno-economic and environmental assessment of a hybrid renewable energy system using multi-objective genetic algorithm: A case study for remote Island in Bangladesh. Energy Convers Manag 2021;230:113823. https://doi.org/10.1016/j.enconman.2020.113823.

[8] Afrinaldi F, Zhang H-C. A fuzzy logic based aggregation method for life cycle impact assessment. J Clean Prod 2014;67:159–72. https://doi.org/10.1016/j.jclepro.2013.12.010.

[9] Goedkoop M, Spriensma R. The Eco-Indicator 99: A Damage Oriented Method for Life Cycle Impact Assessment 2001.

[10] Le Roux D, Lalau Y, Rebouillat B, Neveu P, Olivès R. Thermocline thermal energy storage optimisation combining exergy and life cycle assessment. Energy Convers Manag 2021;248:114787. https://doi.org/10.1016/j.enconman.2021.114787.

[11] Hafizan C, Noor ZZ, Abba AH, Hussein N. An alternative aggregation method for a life cycle impact assessment using an analytical hierarchy process. J Clean Prod 2016;112:3244–55. https://doi.org/10.1016/j.jclepro.2015.09.140.

[12] Agarski B, Budak I, Vukelic D, Hodolic J. Fuzzy multi-criteria-based impact category weighting in life cycle assessment. J Clean Prod 2016;112:3256–66. https://doi.org/10.1016/j.jclepro.2015.09.077.

[13] Sohn J, Bisquert P, Buche P, Hecham A, Kalbar PP, Goldstein B, et al. Argumentation Corrected Context Weighting-Life Cycle Assessment: A Practical Method of Including Stakeholder Perspectives in Multi-Criteria Decision Support for LCA. Sustainability 2020;12:2170. https://doi.org/10.3390/su12062170.

[14] Azapagic A, Clift R. The application of life cycle assessment to process optimisation. vol. 23. 1999.

[15] Ayres RU. Commentary on the utility of the ecological footprint concept. vol. 32. 2000.

[16] Zacharopoulos L, Thonemann N, Dumeier M, Geldermann J. Environmental optimization of the charge of battery electric vehicles. Appl Energy 2023;329:120259. https://doi.org/10.1016/j.apenergy.2022.120259.

[17] Arfan M, Eriksson O, Wang Z, Soam S. Life cycle assessment and life cycle costing of hydrogen production from biowaste and biomass in Sweden. Energy Convers Manag 2023;291:117262. https://doi.org/10.1016/j.enconman.2023.117262.

[18] Lee D, Nam H, Won Seo M, Hoon Lee S, Tokmurzin D, Wang S, et al. Recent progress in the catalytic thermochemical conversion process of biomass for biofuels. Chem Eng J 2022;447:137501. https://doi.org/10.1016/j.cej.2022.137501.

[19] Yong KJ, Wu TY. Second-generation bioenergy from oilseed crop residues: Recent technologies, techno-economic assessments and policies. Energy Convers Manag 2022;267:115869. https://doi.org/10.1016/j.enconman.2022.115869.

[20] Moodley P, Gueguim Kana EB. Development of a steam or microwave-assisted sequential salt-alkali pretreatment for lignocellulosic waste: Effect on delignification and enzymatic hydrolysis. Energy Convers Manag 2017;148:801–8. https://doi.org/10.1016/j.enconman.2017.06.056.

[21] Rangaiah G, Sharma S, Sreepathi BK. Multi-objective optimization for the design and operation of energy efficient chemical processes and power generation. Curr Opin Chem Eng 2015;10:49–62. https://doi.org/10.1016/j.coche.2015.08.006.

[22] Jeon PR, Moon J-H, Ogunsola NO, Lee SH, Ling JLJ, You S, et al. Recent advances and future prospects of thermochemical biofuel conversion processes with machine learning. Chem Eng J 2023;471:144503. https://doi.org/10.1016/j.cej.2023.144503.

[23] Aspen Technology I. ASPEN Plus V12.1 2021.

[24] Humbird D, Davis R, Tao L, Kinchin C, Hsu D, Aden A, et al. Process Design and Economics for Biochemical Conversion of Lignocellulosic Biomass to Ethanol: Dilute-Acid Pretreatment and Enzymatic Hydrolysis of Corn Stover. Golden, CO (United States): 2011. https://doi.org/10.2172/1013269.

[25] Shi S, Guan W, Kang L, Lee YY. Reaction kinetic model of dilute acid-catalyzed hemicellulose hydrolysis of corn stover under high-solid conditions. Ind Eng Chem Res 2017;56:10990–7. https://doi.org/10.1021/acs.iecr.7b01768.

[26] Lavarack BP, Griffin GJ, Rodman D. The acid hydrolysis of sugarcane bagasse hemicellulose to produce xylose, arabinose, glucose and other products. Biom Bioen 2002;23:367–80. https://doi.org/10.1016/S0961-9534(02)00066-1.

[27] Humbird D, Aden A. Biochemical Production of Ethanol from Corn Stover: 2008 State of Technology Model. Technical Report NREL/TP-510-46214: 2009.

[28] Kadam KL, Rydholm EC, McMillan JD. Development and Validation of a Kinetic Model for Enzymatic Saccharification of Lignocellulosic Biomass. Biotechnol Prog 2004;20:698–705. https://doi.org/10.1021/bp034316x.

[29] Leksawasdi N, Joachimsthal EL, Rogers PL. Mathematical modelling of ethanol production from glucose/xylose mixtures by recombinant Zymomonas mobilis. Biotechnol Lett 2001;23:1087–93. https://doi.org/10.1023/A:1010599530577.

[30] Moomaw W, Burgherr P, Heath G, Lenzen M, Nyboer J, Verbruggen A. Annex II: Methodology. . In IPCC Special Report on Renewable Energy Sources and Climate Change Mitigation [O. Edenhofer, R. Pichs-Madruga, Y. Sokona, K. Seyboth, P. Matschoss, S. Kadner, T. Zwickel, P. Eickemeier, G. Hansen, S.

Schlömer, C. von Stechow (eds)]. Cambridge, United Kingdom and New York, NY, USA: 2011.

[31] 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;21:1218–30. https://doi.org/10.1007/s11367-016-1087-8.

[32] Huijbregts MAJ, Steinmann ZJN, Elshout PMF, Stam G, Verones F, Vieira MDM, et al. ReCiPe2016. A harmonized life cycle impact assessment method at midpoint and endpoint level. Report I: Characterization. RIVM Report 2016-0104. Bilthoven: 2017.

[33] Fozer D, Nimmegeers P, Toth AJ, Varbanov PS, Klemeš JJ, Mizsey P, et al. Hybrid Prediction-Driven High-Throughput Sustainability Screening for Advancing Waste-to-Dimethyl Ether Valorization. Environ Sci Technol 2023;57:13449–62. https://doi.org/10.1021/acs.est.3c01892.

[34] Mitchell DP. Spectrally optimal sampling for distribution ray tracing. Proceedings of the 18th annual conference on Computer graphics and interactive techniques - SIGGRAPH '91, New York, New York, USA: ACM Press; 1991, p. 157–64. https://doi.org/10.1145/122718.122736.

[35] Kamath C. Intelligent sampling for surrogate modeling, hyperparameter optimization, and data analysis. Machine Learning with Applications 2022;9:100373. https://doi.org/10.1016/j.mlwa.2022.100373.

[36] Berry MJA, Linoff GS. Data Mining Techniques For Marketing, Sales, and Customer Relationship Management. 2nd ed. Wiley Publishing Inc.; 2004.

[37] Vasilakou K, Nimmegeers P, Thomassen G, Billen P, Van Passel S. Assessing the future of second-generation bioethanol by 2030 – A techno-economic assessment integrating technology learning curves. Appl Energy 2023;344:121263. https://doi.org/10.1016/j.apenergy.2023.121263.

[38] Kalbar PP, Birkved M, Nygaard SE, Hauschild M. Weighting and Aggregation in Life Cycle Assessment: Do Present Aggregated Single Scores Provide Correct Decision Support? J Ind Ecol 2017;21:1591–600. https://doi.org/10.1111/jiec.12520.

[39] Nimmegeers P, Vallerio M, Telen D, Van Impe J, Logist F. Interactive Multi-objective Dynamic Optimization of Bioreactors under Parametric Uncertainty. Chemie Ingenieur Technik 2019;91:349–62. https://doi.org/10.1002/cite.201800082.

[40] Nimmegeers P, Telen D, Logist F, Impe J Van. Dynamic optimization of biological networks under parametric uncertainty. BMC Syst Biol 2016;10:86. https://doi.org/10.1186/s12918-016-0328-6.

[41] Mores W, Nimmegeers P, Hashem I, Bhonsale SS, Van Impe JFM. Multi-objective optimization under parametric uncertainty: A Pareto ellipsoids-based algorithm. Comput Chem Eng 2023;169:108099. https://doi.org/10.1016/j.compchemeng.2022.108099.