



HAL
open science

Integration of environmental analysis in the assessments of hybrid energy systems

Hemant Sharma

► **To cite this version:**

Hemant Sharma. Integration of environmental analysis in the assessments of hybrid energy systems. Eco-conception. Université Grenoble Alpes [2020-..], 2021. English. NNT : 2021GRALI045 . tel-04055302

HAL Id: tel-04055302

<https://theses.hal.science/tel-04055302>

Submitted on 2 Apr 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

Pour obtenir le grade de

DOCTEUR DE LA COMMUNAUTÉ UNIVERSITÉ GRENOBLE ALPES

Spécialité : **Génie industriel**

Arrêté ministériel : 25 mai 2016

Présentée par

Hemant SHARMA

Thèse dirigée par **Peggy ZWOLINSKI**
et co-encadrée par **Guillaume MANDIL, Élise MONNIER et Emmanuelle COR**

préparée au sein du **Laboratoire CEA-Liten**
dans **l'École Doctorale I-MEP2**

Integration of environmental analysis in the assessments of hybrid energy systems

Thèse soutenue publiquement le **26 Mars 2021**,
devant le jury composé de :

Monsieur Frédéric WURTZ

Directeur de Recherche au CNRS, Président

Madame Catherine AZZARO PANTEL

Professeure, INP-ENSIACET, Rapporteur

Monsieur Nicolas PERRY

Professeur I2M, ENSAM, Rapporteur

Madame Peggy ZWOLINSKI

Professeure, Grenoble INP, Directrice de thèse

Monsieur Guillaume MANDIL

Maître de Conférences, UGA, Co-Encadrant de thèse

Madame Élise MONNIER

Ingénieure, CEA-Liten, Co-Encadrante de thèse

Madame Emmanuelle COR

Ingénieure de Recherche, CEA-Liten, Co-Encadrante de thèse



"Dedicated to my Grandparents"

Contents

Table of contents	i
General introduction	1
1 Establishment of the thesis context	4
1.1 Modelling context: Emerging technologies in hybrid energy simulations	5
1.1.1 Activities conducted in the laboratory	5
1.1.2 System level analysis - hybrid energy systems	5
1.1.3 Hybrid energy systems: a pilot scale implementation of emerging technologies	6
1.1.4 Energy simulation tool used in this thesis - Odyssey	6
1.1.5 Example: case study of hydrogen production	10
1.2 Taking environmental impact into account	14
1.2.1 Need for environmental assessment	14
1.2.2 Relevance of life cycle assessment (LCA)	15
1.2.3 Laboratory engineers and life cycle assessment	16
1.3 Problems in the status-quo with respect to the inclusion of environmental impacts	17
1.3.1 Problem 1: Inadequate environmental impact estimation	17
1.3.2 Problem 2: Exclusion of LCA indicators in the optimisation process	18
1.4 Goals of the thesis	21
1.4.1 Integrated LCA modelling along with hybrid energy simulations	21
1.4.2 Inclusion of LCA indicators in multi-objective optimisation	22
1.5 Research Approach	22
2 Positioning the thesis in the existing literature	24
2.1 Environmental information in the hybrid energy simulations	25
2.1.1 Summary: Insufficient environmental information	26
2.2 LCA computational models	27
2.2.1 Computational structure of LCA	27
2.2.2 Search Method	28
2.2.3 Selection criteria	29
2.2.4 Results	29
2.2.5 Summary: Parameterized LCA models are selected but some issues remain unsolved	37
2.3 Inclusion of LCA indicators in optimisation	38
2.3.1 Problems with the currently used genetic algorithm SPEA2	38
2.3.2 Use of genetic algorithm NSGA-III in this thesis	39
2.3.3 Optimisation overview in life cycle assessment	43
2.3.4 Objective reduction techniques in life cycle assessment	44
2.3.5 Summary: A possible combination of NSGA-III with objective reduction?	47
2.4 Synthesis of research proposals	48

3	Establishment of parameterized life cycle assessment models for hybrid energy simulations	51
3.1	Introduction	53
3.1.1	Contribution to the existing literature	53
3.1.2	Framework development approach	53
3.1.3	Relevance of the framework	54
3.2	Framework to establish parameterized LCA models for hybrid energy simulations	54
3.2.1	Handling inputs from energy simulations	55
3.2.2	LCA modelling approach	55
3.2.3	Dealing with variabilities and uncertainties	57
3.2.4	Model assembly and parameter definition	61
3.2.5	GSA using Derivative-based Global Sensitivity Measure	63
3.3	Case study: PEM water electrolysis	66
3.3.1	Handling outputs from the energy simulation	67
3.3.2	LCA modelling approach	68
3.3.3	Addressing variabilities and uncertainties	69
3.3.4	Model Assembly	70
3.3.5	Parameter definition for GSA	71
3.4	Case study: Results	72
3.4.1	Sensitivity & Importance indices	72
3.4.2	Improving the influential aspects for the scope	73
3.4.3	Results for scope 2030 and comparison with ideal case & literature values	74
3.4.4	When influential parameters cannot be fixed	75
3.5	Conclusions	76
4	Objective reduction of LCA indicators using a visual survey	79
4.1	Introduction	80
4.1.1	Contribution to literature and relevance	80
4.1.2	Type of LCA indicators: midpoints and endpoints	80
4.1.3	Impact assessment method selected in this thesis	83
4.2	Indicator selection based on a visual survey	83
4.2.1	Example: Design of a bike	84
4.2.2	Design of a bike: comparison with PCA	87
4.2.3	Limitations	89
4.3	Recommendations on its use	90
4.4	Conclusions	90
5	Including environmental indicators in the multi-objective search of solutions	92
5.1	Relevance and contribution to the literature	93
5.2	Hydrogen production case study	94
5.2.1	Inputs to the integrated model	95
5.2.2	Computational structure	97
5.3	Performance of NSGA-III with reduced indicators	98
5.3.1	Hypervolume indicator	99
5.3.2	Indicator reduction using visual survey	100
5.3.3	Indicator reduction using PCA	102
5.3.4	Reference point specification	104
5.3.5	Results with NSGA-III and reduced indicators	105
5.3.6	Discussion	110
5.4	Visualization of solutions for decision-making	112
5.4.1	Proposed procedure	112

5.4.2	Application to the obtained results	113
5.4.3	Comparison with the status-quo	117
5.5	Conclusions	118
General conclusions		122
Perspectives		127
Abbreviations · Publications		I
Bibliography		III
List of figures		XXII
List of tables		XXVI
A	Annex 1 : Supplementary data	XXIX
B	Annex 2 : Example output of parameterised wind model	XXXVII
C	Annex 3 : Supplementary figures from chapter 5	XLI

General introduction

Since the 1950s, negative effects due to rise in global greenhouse gas emissions emanating from human activities have been observed. Atmospheric and ocean temperatures have risen, ice caps around the world are melting. Climate change is already causing changes in extreme weather events leading to a rise in the intensity and frequency of floods, hurricanes, droughts, heatwaves and wildfires [IPCC2014].

Deleterious effects of human activities on our ecosystems have also been observed. These ecosystems provide the basic support for human life by maintaining earth habitable. For example, clean air, water and food essential for human survival are strongly dependent on natural systems [Brondizio2019]. However they are threatened due to rapidly growing human needs for food, energy, water and materials.

We are entering a 'mass extinction' of species which is threatening the biodiversity and consequently the balance of the ecosystems [Ceballos2015]. For example, extinction rates of various animal groups (vertebrates, birds, reptiles, etc.) are predicted to be 8-100 times higher since 1500s. Around 40 percent of amphibians as well as 33 percent of corals and marine mammals are threatened with extinction [Brondizio2019].

Energy and electricity generation as a sector is the largest contributor to the greenhouse gas emissions. This can be verified in figure 1, where global greenhouse gas emissions from 2016 are displayed. This is mainly due to the use of fossil fuel based technologies: coal, gas and oil.

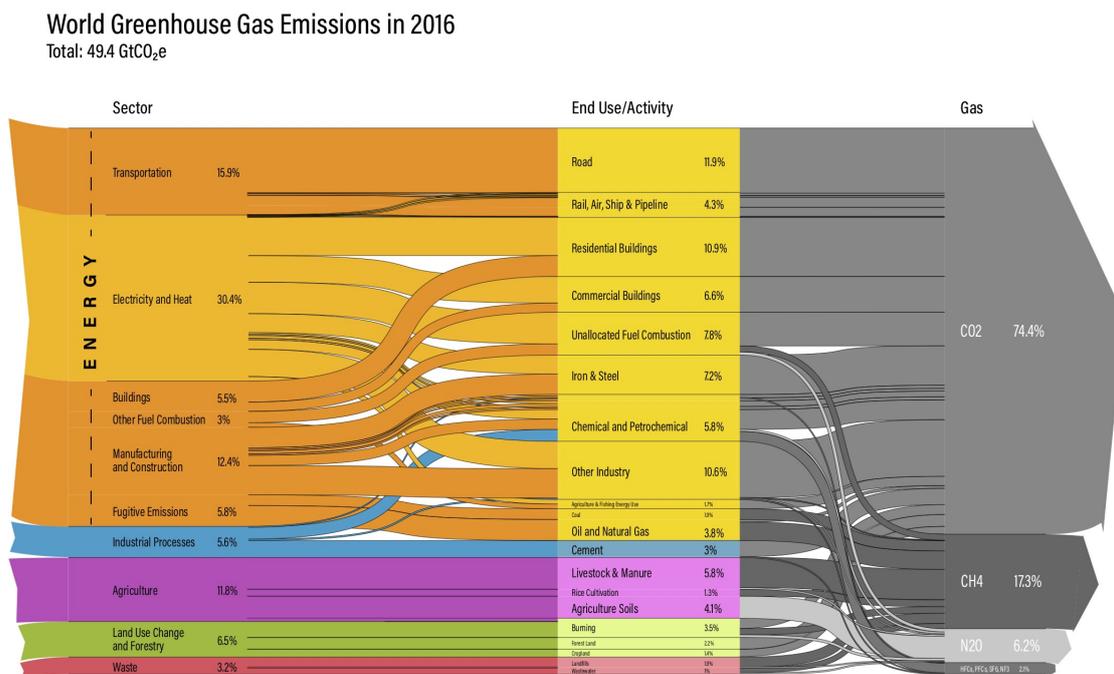


Figure 1: Breakdown of world greenhouse gas emissions in 2016, updated from [Herzog2009].

Consequently, new energy technologies have been proposed to keep greenhouse gas emissions in check. These technologies e.g. wind turbines and solar PV, could have lower carbon emissions by an order of magnitude as compared to their fossil counterparts. Nonetheless, it does not mean that they also have lower negative effects on human health, natural resources or ecosystems with respect to other environmental indicators.

The technologies entering the market should thus also be assessed from an environmental criteria over its entire life cycle and with a variety of different indicators such as acidification, ecotoxicity, etc. This would enable us to identify and limit their impacts before the possible commercialisation of such technologies. Assessing them at a developing stage would also avoid large scale environmental problems in the future. Integrating this criteria into assessments would facilitate optimal decision-making regarding these systems. Furthermore, it will also enable planning these facilities such that their environmental impacts are minimised.

Hence, the aim of this thesis is to propose an approach to integrate environmental information in the assessments of hybrid energy systems. These systems are often the stage where emerging energy technologies are demonstrated and their performance in real world conditions is quantified. More specifically, two main problems are identified that prevent the inclusion of environmental information. First, environmental information was not calculated with the corresponding changes in energy simulations. Second, environmental indicators were not included in the multi-objective optimisation. Research proposals were introduced for addressing these two problems. A case study involving hydrogen production from electrolysis of water is used to illustrate and validate the concepts introduced in this thesis.

1

Establishment of the thesis context

Summary: This chapter first presents the context of emerging technologies in hybrid energy simulations, in which this thesis was written. The need of including environmental impacts at this stage of analysis is then described. To satisfy this need, two main problems arise in the status-quo. They are elaborated using relevant arguments and examples. Then, to solve these problems, two main goals of the the thesis are presented.

Contents

1.1	Modelling context: Emerging technologies in hybrid energy simulations . .	5
1.1.1	Activities conducted in the laboratory	5
1.1.2	System level analysis - hybrid energy systems	5
1.1.3	Hybrid energy systems: a pilot scale implementation of emerging technologies	6
1.1.4	Energy simulation tool used in this thesis - Odyssey	6
1.1.5	Example: case study of hydrogen production	10
1.2	Taking environmental impact into account	14
1.2.1	Need for environmental assessment	14
1.2.2	Relevance of life cycle assessment (LCA)	15
1.2.3	Laboratory engineers and life cycle assessment	16
1.3	Problems in the status-quo with respect to the inclusion of environmental impacts	17
1.3.1	Problem 1: Inadequate environmental impact estimation	17
1.3.2	Problem 2: Exclusion of LCA indicators in the optimisation process . .	18
1.4	Goals of the thesis	21
1.4.1	Integrated LCA modelling along with hybrid energy simulations	21
1.4.2	Inclusion of LCA indicators in multi-objective optimisation	22
1.5	Research Approach	22

1.1 Modelling context: Emerging technologies in hybrid energy simulations

The aim of this section is to present the background of the thesis, more specifically on the hybrid energy simulations. First, the activities conducted in the laboratory are presented. Then, the positioning of this thesis at the system level is elaborated and the interest in analysing technologies at this level is also described. Energy simulation software used in this thesis, Odyssey, is then presented with the help of an example.

These results were obtained from discussion with the laboratory personnel and through the utilisation of the software tools in the lab. The laboratory personnel are also the direct or indirect users of the energy simulation tools. Hence, they are referred to as 'Users' hereafter. Discussions were held with them using open-ended questionnaires.

1.1.1 Activities conducted in the laboratory

The thesis was conducted in the laboratory of energy systems and demonstrations (LSED) at CEA-Liten (French commission for atomic and alternative energies). The laboratory LSED works on technical, economic and environmental evaluation of energy and district heating systems at three levels. The three levels being: component, system and territory. At the component level, detailed analysis of only of a technology is done with limited interactions with other components. At the system level as the name suggests, multiple technologies/components are modelled as close to the real conditions as possible. Finally, the territory level is analysis energy systems at a regional scale. An example can be assessing the heating or the electricity grid of a city. Another important point to note is that the laboratory mainly deals with emerging energy technologies and its corresponding solutions. The technology readiness levels (TRL) dealt with is around 3-8. The aim is to evaluate these systems and provide recommendations on investment and R&D

1.1.2 System level analysis - hybrid energy systems

This thesis is primarily situated at the system level, even though certain principles could also be applied at other levels. This 'level' of system analysis is also called hybrid energy simulation or local/community-scale in the literature [Connolly2010, Sinha2014]. Hybrid energy system is defined as the one which incorporates multiple energy generation and storage options, along with the load [Nazari-Heris2018]. Another salient feature that is often associated with hybrid energy systems is the presence of renewable energy sources with or without the conventional generation [Connolly2010]. Furthermore, the energy demand or generation can be in different forms: heating, electricity or even fuels such as hydrogen. For example, a hybrid energy system can be a combination of photovoltaics (PV), wind turbines, grid-electricity, batteries, fuel cells, hydrogen source and so on. These systems can be operated with a connection to central electricity grid/heating network, or without it, in stand-alone mode.

During the exchanges with the laboratory personnel, they were asked to provide examples of some their past projects. Below is the list of technologies mentioned by the users when describing examples of projects. Hence, it is a non-exhaustive list.

- Photovoltaics
- Wind Turbines
- Hydrogen storage
- Batteries: Li-ion and lead-acid
- Methanation
- Fuel cells and Electrolysers: Alkaline, PEM and SOEC
- Gasification technologies
- Biogas from anaerobic digestion
- Electricity/gas network

It can be seen that some of technologies above are not yet fully commercialised and are still in the 'emerging' stage. The technologies can also be referred to as components of hybrid energy systems.

1.1.3 Hybrid energy systems: a pilot scale implementation of emerging technologies

Hybrid energy systems are widely used for pilot scale-demonstration of emerging technologies. This is a crucial step in their development before their eventual commercialisation, since they are evaluated in an operational context i.e. in real world conditions. It is thus also the first time when a particular technology can be legitimately compared to its alternatives. Analysis at this level provides important insights with respect to operation of a technology, interaction with other components and influence of on-ground conditions. Consequently, it facilitates planning of these systems, identification of key issues and quantification of future performance. Thus the projects usually have one of the following objectives: feasibility analysis, optimisation of operation or dimensioning.

In order to conduct such assessments, reliable simulations are necessary. According to User2 "*(Simulations) enable us to see how technologies perform with load, local conditions, constraints, etc.*". Another User11 further adds "*It is essential to use simulations for prospective studies...to evaluate energy scenarios or specific designs in order to assess their validity and performance*".

Furthermore, these simulations, are often quite complex due to the following:

- Variability and mismatch in the renewable resources and the demand
- Multiple technological options possible to satisfy the same objective
- Interactions between different energy carriers: electricity, hydrogen, heating, etc.
- Trade-offs between the different performance indicators

To deal with these complexities, understand their behaviour and effectively design them, 'dynamic' simulation are necessary. Dynamic simulation means the simulations taking into account the time dependent variations in the functioning of the system. It is particularly relevant in this context owing to the intermittent nature of energy components in question. User8 notes this importance "*Clients are interested in dynamic behaviour of their energy systems*".

1.1.4 Energy simulation tool used in this thesis - Odyssey

In this thesis, the software 'Odyssey' developed by CEA is used for the energy simulations [Guinot2013]. It is primarily used internally by the laboratory members and simultaneously developed on a project-by-project basis. It can work at a system level up to a territory or regional level, integrating renewable sources and storage. It is principally used for techno-economic feasibility analysis or dimensioning. Any other hybrid energy software could also have been used instead. The software is briefly described below and illustrated in figure 1.1.

Inputs

The inputs are obtained from clients or partners (in case of collaborative projects). In some cases, inputs are searched or estimated by the users themselves along with a discussion with stakeholders. They can be classified in the following types:

Time-series. It enables entering variable values for the models over a period of one year. Time-step can vary from 1s to 1 day. It may consist of variable renewable sources, energy spot prices, energy demand, etc.

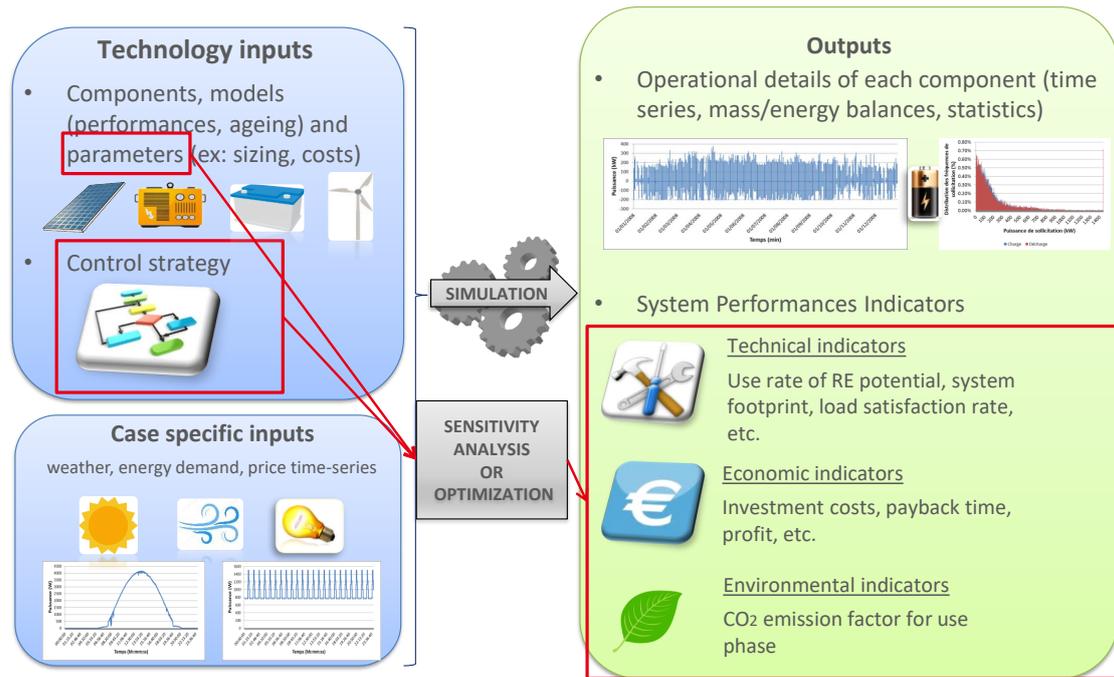


Figure 1.1: An overview of the energy simulation software used in this thesis - Odyssey.

Model parameters. They consist of different parameters that influence the system modelling. They may consist of technical or economic parameters either at a level of particular technology components or at a level of system. It is possible to define models with different levels of detail. For example, including efficiency degradation in batteries and taking into account the replacement during the project lifetime.

Control strategy. Operation rules need to be specified which essentially dictate prioritisation of components in case of conflicts. For example, use of local vs. grid electricity.

Functions

The users thus choose a system architecture i.e. a combination of technology components, define model parameters, enter time-series and finally define a pertinent operation strategy. Time-series can be obtained from historically recorded data, online databases or even generated artificially.

Simulations. Using the above inputs, simulation of a single system configuration can be obtained. A system configuration means particular size of each component and a specific operation strategy. The software then calculates component behaviour at each time-step for a duration of one year. The results are then extrapolated for a duration of the project lifetime. Consequently, this enables the calculation of performance indicators. Technical indicators typically measure the energy produced or satisfied or un-utilised. On the other hand, economic indicators deal with the associated costs of energy or system.

Then, further investigation of system parameters or operation through sensitivity analysis or multi-criteria optimisation is also possible. Since it strongly depends on the project and the allocated resources, these two steps are thus optional. Techniques compatible with 'black-box' models are utilised since the simulations cannot be mathematically formulated in this software. That is, the simulations performed cannot be represented as a mathematical equation. In brief, these techniques initialise different system configurations and then analyse the

performance indicators to find influential parameters or to search optimal solutions. Both of these techniques will be described in detail in the coming chapters.

Sensitivity analysis. There are often many uncertainties while modelling a system. In this step, the uncertainties are specified, followed by the quantification of their influence on the final results. Local or global sensitivity analysis are currently possible. In the former case, value of a parameter is varied around a reference point keeping other values the same to get a possible range of outputs. While global sensitivity analysis methods (GSA e.g. Sobol indices or Morris method) are also applied to find the most influential parameters from a statistical distribution of outputs due to the uncertain parameters [Nadal2019].

Multi-criteria optimisation is used to search system configurations with respect to max 3 criteria simultaneously using the evolutionary algorithm SPEA2 [Zitzler2001]. Hence, it can be useful for optimising operation of a system or for dimensioning a system. The criteria or the objectives are nothing but the performance indicators of a system, e.g. cost, energy demand satisfied. The multi-objective decision-making is especially useful in this context to find solutions with trade-offs between the different objectives. Example, lower cost might mean that the energy demand is not completely satisfied and vice-versa. Then a decision can be made on the compromise that best suits the desired requirements. Multi-criteria optimisation is also combined with global sensitivity analysis to facilitate design under uncertainties [Nadal2020].

Optimisation with SPEA2

Since multi-criteria optimisation forms a major part of this thesis, the SPEA2 algorithm is briefly described in this section. Before that however, the jargons used in this field are introduced below while placing them in the context of this thesis.

Individual is a particular set of variables for which optimal values are sought. There can be thus multiple variables to be optimised at the same time relating to parameters such as size or control strategy. Optimisation variables could have discrete or continuous values. In this thesis, the focus is on dimensioning the system. Thus, an individual in this document is a system configuration with a particular size of each energy component in the system.

Optimisation objective or optimisation criteria is the one with respect to which the set of variables are to be optimised. It can be a single objective or multiple for a multi-objective optimisation. When the optimisation objectives are 2-3, it is called multi-objective optimisation. Whereas when they increase more than 3, it is often referred to as many-objective optimisation in the literature [Bechikh2016]. In this thesis, these are the performance indicators of a system configuration that can be technical, economic or environmental. Prior to this thesis, the number of criteria for simultaneous optimisation in Odyssey was maximum 3.

Population is a collection of individuals. Furthermore, parents is the population from which offsprings are obtained, usually by applying mate and mutation operators.

Mate and mutation operators can be seen in figure 1.2. These two are randomly applied to individuals. Mating also known as crossover, combines optimisation variables of two individuals while mutation changes value of a variable of an individual.

Fitness is the performance indicator of the individual on the basis of which it is evaluated. Often it is the same as the objective value of the individual. Then depending if the problem is about minimisation or maximisation of the objective, fitness of an individual is better, worse or comparable with the other. Another difference is that some algorithms (like SPEA2) might use relative fitness values for computations instead of absolute objective values.

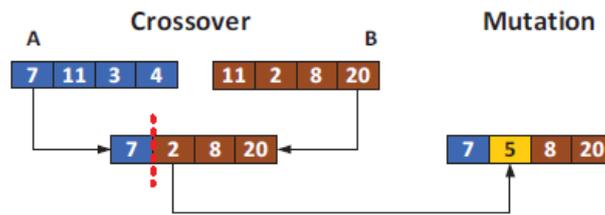


Figure 1.2: Crossover and mutation operations, illustration from [Perez Gallardo2013].

Domination An individual X is said to dominate another individual Y when the fitness of X is better than Y in at least one objective while the fitnesses of X in all other objectives is not worse than Y. On the other hand, if fitness of X is worse than Y in at least one objective, and the same is true for fitness of Y with respect to X, then the two solutions are said to be non-dominated.

Pareto front is made up of non-dominated individuals. The main property of this curve is such that value of an objective cannot be improved without degrading a value of another objective. Thus, it offers the trade-offs between the objectives simultaneously optimised. For a mono-criteria optimisation, this is only a single point, better than all other individuals. For a n-objective optimisation, it consists of a curve of non-dominated individuals in n-dimensions.

Diversity is the measure of *different* individuals in the population. There is no single measure for diversity as it may refer to values of fitness or individuals types. In this thesis, diversity is referred in terms of different fitness values. It becomes more and more important as the number of optimisation objectives increase. As they increase, the pareto front increases exponentially. Since in an algorithm, a finite number of solutions can be searched, it is important to search solutions that represent different (or preferential) parts of the pareto front. Lets say three individuals have the following objective values as (demand satisfaction, cost and carbon emissions): (1%, 11 Euros, 6 kg CO₂-eq), (1.5%, 10 Euros, 6.1 kg CO₂-eq) and (8%, 12 Euros, 2 kg CO₂-eq). If two solutions had to be selected between the three, it is better to retain the last one in any set since it offers a different perspective of higher unsatisfied demand but lower carbon emissions.

Now that the essential concepts have been discussed, SPEA2 can be described, a simplified version is illustrated in figure 1.3. In the first step, a random population is initialised. All the non-dominated individuals (based on their objective values) are directly copied to the archive. If the archive size (S i.e. no. of individuals in the archive) increases more than a pre-defined value, individuals maintaining the diversity are preferentially retained. Fitness values are then assigned to each individual in the archive and the population. As mentioned earlier, these are relative fitness values depending on the number of individuals dominated by it. The fitness assignment equations are not described here. Then, all the non-dominated individuals in the population are copied to a set (size N). If $N > S$, individuals are deleted using the diversity criteria with more diverse individuals retained. If $N < S$, next best dominated solutions from the population are added. If N equals S, algorithm proceeds with the next step.

If the termination criteria (or the stopping condition) is not satisfied, N solutions are copied to the archive. Individuals are selected from N for the next generation using binary tournaments. In such a selection, one individual is selected from randomly grouped 2 individuals, based on the fitness values. Mutation and crossover operators are then applied to form the population of size P for the next generation. The loop thus continues until the termination criteria is satisfied. When

it is, the non-dominated solutions in N are presented as the final solutions. An example of the termination criteria could be maximum number of generations reached.

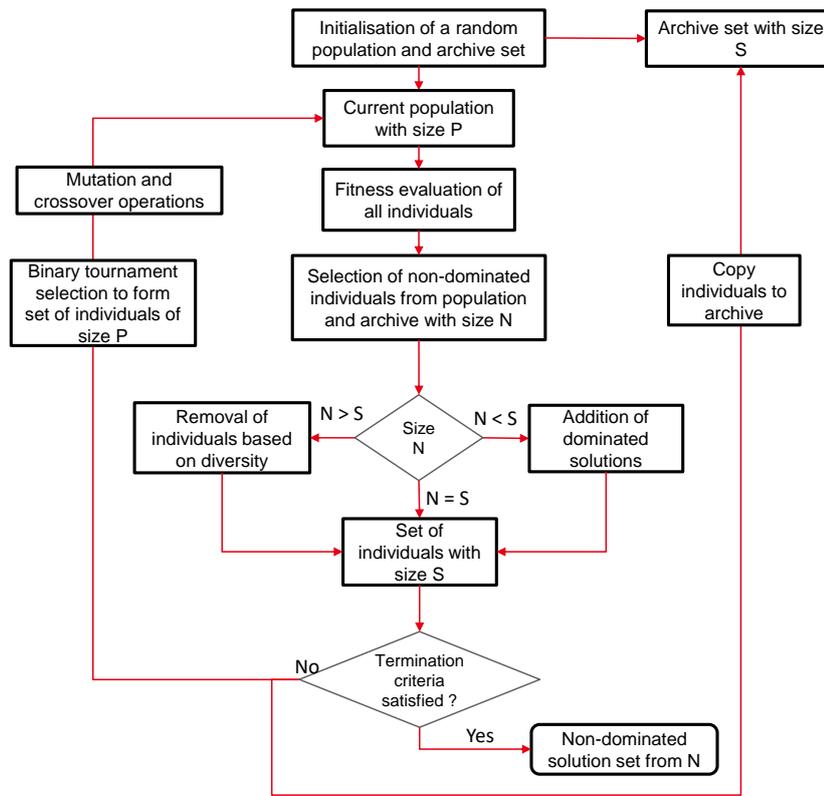


Figure 1.3: SPEA2 operating scheme as described in [Zitzler2001].

1.1.5 Example: case study of hydrogen production

To facilitate the communication of the concepts introduced in the previous sub-section, an example case study implemented in Odyssey is presented here. This case study can be seen as a reference of the current state of analysis possible for the hybrid energy systems. The same case will be modified in the future chapters for the demonstration of the proposals.

Goal

The location is Marseille, France where a variable hydrogen load is satisfied using water electrolysis. The electricity is supplied by the grid and there is a H_2 storage. The aim is to minimise the hydrogen production cost while satisfying the hydrogen demand as much as possible. Thus, to dimension the following components: Electrolyser nominal power (ELY) and H_2 storage size (Tank H_2). The case study as seen in the software interface is illustrated in figure 1.4.

Inputs

The input parameters are kept simple and default values are used where necessary. Hence, the results are only for illustration purposes. Electrolyser efficiency is 56%, based on the lower heating value of hydrogen. Cost of electrolyser and the storage varies linearly according to the size. For the dynamic simulation, time series of hydrogen demand and electricity spot prices, for every five minutes are available for the entire year and is illustrated for the initial nine days in figure 1.5.

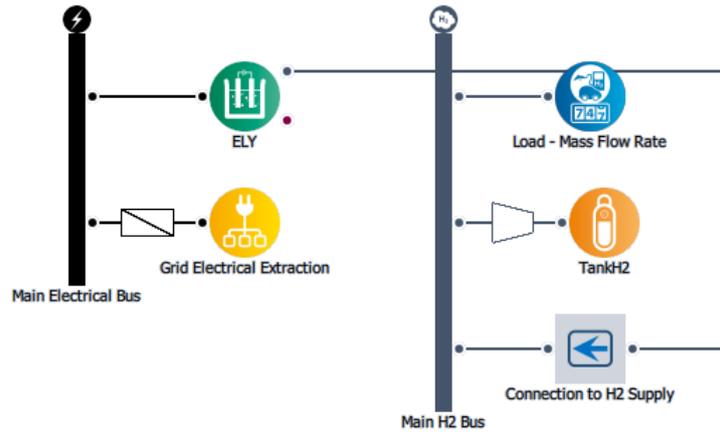


Figure 1.4: Case study of hydrogen production from PEM water electrolysis.

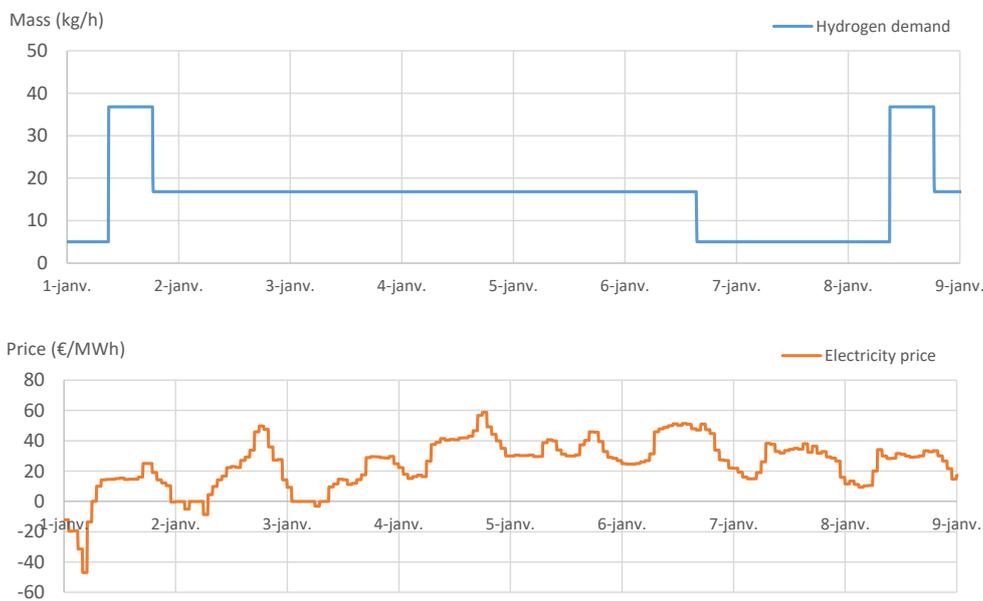


Figure 1.5: Inputs time series: hydrogen demand (kg/h) and electricity price (€/MWh).

There are several operational strategies available, which essentially is a way of instructing the software how to operate the system when there are multiple ways of operation are possible. In this case, electricity price is unusually low even going below zero on multiple occasions due to excess electricity produced by a nearby source. Therefore, we use a strategy to force electrolyser to use this electricity when the price is below 50 €/MWh.

The goal is to dimension the electrolyser size and the tanks, with respect to the hydrogen production cost (€/kg) while satisfying the maximum hydrogen demand. The latter can be measured using the mass based unmet load, which means fraction of the demand (in kg) not satisfied throughout the project lifetime. An approximate range (table 1.1) to size the two components with respect to the above two objectives is given to the optimiser.

Outputs

First the optimisation results are presented. If there was only one objective to be optimised, there would only be one solution. Since there are 2 objectives to be minimised, the resulting solutions

Table 1.1: Summary of inputs to the optimisation

Component	Technical	Economic	Range for dimensioning
Electrolyser	56% conversion efficiency	linear cost model respect to size	0.5 - 1.5 MW
Tanks	30-60 bar pressure range	linear cost model respect to size	1 - 150 m ³

have a trade-off between the two objectives. These solutions are presented as a pareto front in figure 1.6. It can be seen that any solution, cannot be made better off with respect to an objective, without making another objective worse. For example, as seen in the figure, cost of hydrogen for a particular solution cannot be decreased without increasing the unmet load. These solutions represent different sizes of electrolyser and storage. To analyse them further, 4 arbitrary solutions are picked, they are shown in table 1.2. The change in the size of electrolyser-storage and consequently, trade-off between cost-demand satisfaction can be seen clearly.

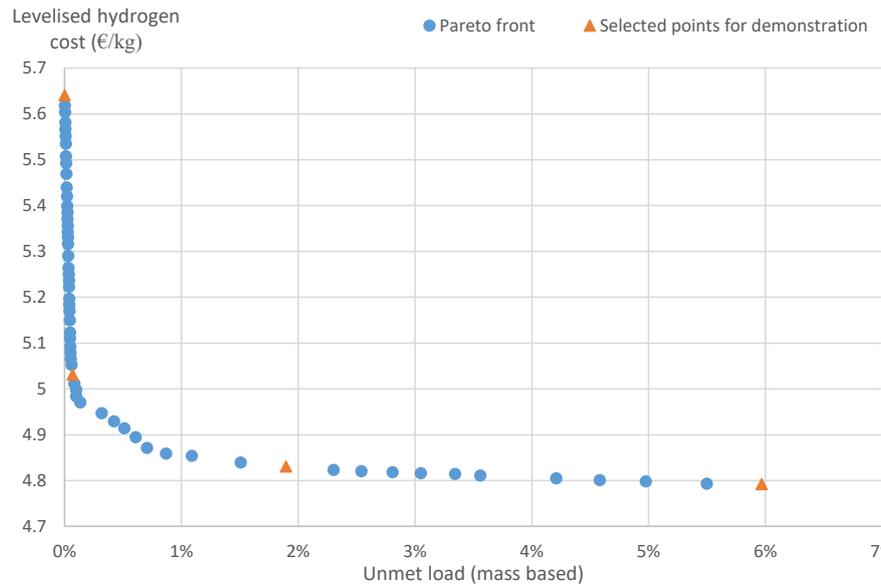


Figure 1.6: Pareto front of the results for the dimensioning problem

Table 1.2: Description of the four selected cases from the pareto front

No.	Unmet load	Levelised hydrogen cost (€/kg)	Electrolyser size (MW)	Storage size (m ³)
1	0.00%	5.64	1.27	21
2	0.07%	5.03	1.05	31
3	1.89%	4.83	0.95	46
4	5.97%	4.79	0.91	21

During the above dimensioning, operational details of the system for each time step were also calculated, in order to calculate the hydrogen cost and unmet load. For example, the electrolyser

operating power and hydrogen state of charge in the storage for case 3 is illustrated in figure 1.7. These representations are useful to see the system behaviour in detail. It also highlights the dynamic nature of results, which are crucial in this context with intermittent energy sources.

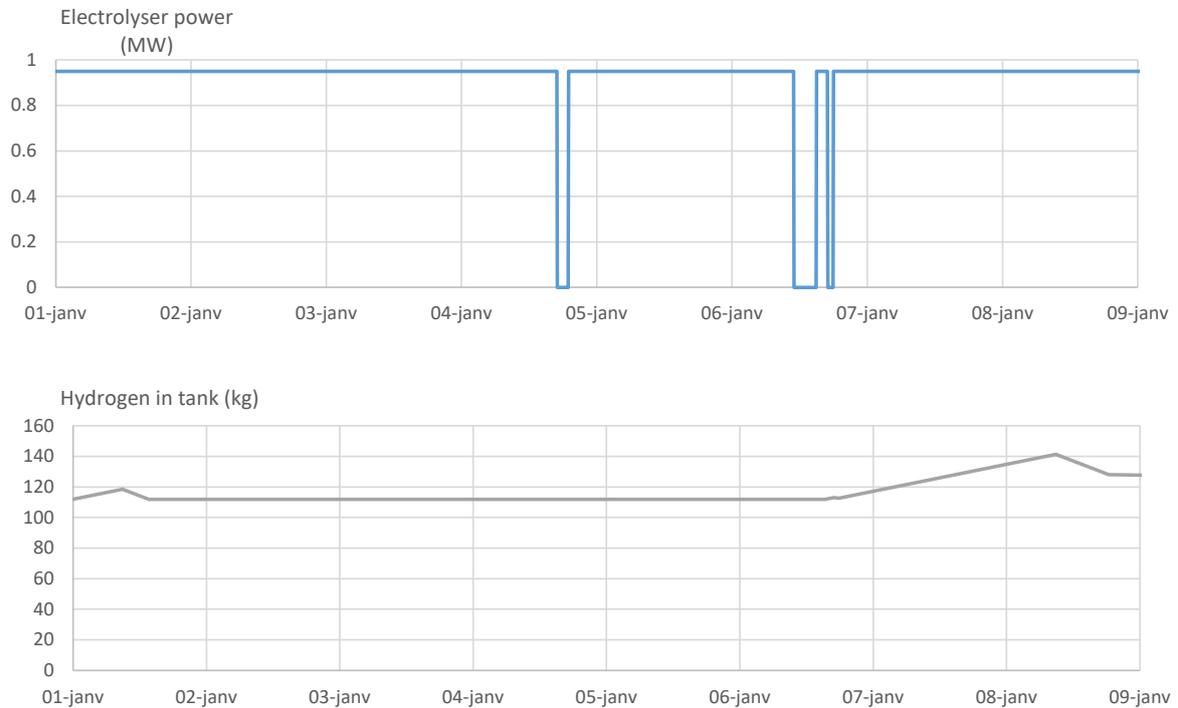


Figure 1.7: Output time series for case no. 3: electrolyser power (MW) and Hydrogen quantity in the tank (kg)

Relevance of multi-criteria optimisation

The results from the pareto front can thus be used for dimensionning a system with respect to the dynamic conditions present. The information about the solutions in the curve is readily available in the format of table 1.2. Here the added value of the multi-criteria optimisation can be seen with respect to the different system configurations possible. For example, if 100% of the hydrogen demand is to be satisfied over the project lifetime, the cost becomes the highest. On the other hand, if some leeway is possible, the cost can be decreased upto 15% while limiting the unsatisfied demand to less than 2%.

The optimisation module provides a systematic way to find the best possible solutions in terms of the relevant criteria. It also helps to comprehend the trade-offs between the criteria for a given case and arrive to a decision accordingly. In the absence of this, users will have to manually simulate different sizes of electrolyser and storage based on some estimates to consequently find their techno-economic performance. Even if that is done, it is quite difficult to know if these solutions lie on or close to the pareto front. In other words, it would be highly probable that solutions exist that are better than the ones manually found in both of the two criteria.

The optimisation and the corresponding analysis is an iterative process. For example, since the above problem is already known, the optimisation bounds in table 1.1 were straightforward to define. However, depending on the case if the optimisation bounds appear in one of the solutions, search using the algorithm has to redone with the modified range. For instance in the case study above, if one of the solutions in the pareto front had contained 1.5 MW (max value of electrolyser)

the search has to be redone after increasing max range. So as to allow the search of systems with these sizes.

Optimisation even though a practical tool, is not required for all projects. In many cases the system is already dimensioned and a further analysis is requested from the lab. Examples can be feasibility of a scaled-up business case or practicality of a system. Detailed performance of each component calculated by the software can be useful to make such an analysis. Sensitivity analysis (local or global) might be used if required.

Relevance of the results

It is possible to analyse each system configuration according to performance of a component at each time step such as figure 1.7. The selected points can be recommended by the users to the decision-makers or even one specific solution if the additional preferences are known. This is also an iterative process involving exchange of information between the users and other stakeholders in a project. Users might receive critiques on their analyses or might obtain future investigation directions. On the other hand, users might demand further details to improve their models on issues such as time-series or a technology behaviour.

A set of important insights provided by this type of analysis is the future research or planning directions. This is useful information for a successful transition of a technology from demonstration to commercial scale. At the moment, relevant information can be obtained for technical and economic criteria. For example, sensitivity of the hydrogen price with respect to the efficiency of the electrolyser v/s grid electricity price could be found in a real world scenario. This information can be useful to guide improvement efforts.

Another example can be planning the electrolyser operation. Assume a hypothetical case when the operation strategy is the optimisation variable and it is found that the hydrogen production is most profitable when electrolyser is operated at 50% power due to electricity price time serie and stack degradation. Hydrogen production from electrolysis can planned in the areas where electricity price is conducive to this or research efforts could be directed to modify/improve this behaviour.

The final decision about implementing one or none of the systems rests with the decision-makers. It could be supplemented by additional analysis which could be done in the same laboratory or by other stakeholders.

1.2 Taking environmental impact into account

In this section, including the environmental impact with the hybrid energy simulations is discussed. First the requirement for this type of analysis is discussed. Then, the relevance of life cycle assessment as a tool for this context along with its description is discussed. The experience of users with LCA and their opinion on different aspects of LCA is presented in the final subsection.

1.2.1 Need for environmental assessment

Heat and electricity generation as a sector contributed the maximum to greenhouse gas emissions in 2016 as was seen in the introduction. Thus there has been a lot of emphasis on emerging, 'low-carbon' technologies in the energy sector. Consequently, during their development, it is important to estimate their environmental impact for a number of reasons such as: comparison with conventional technologies; identify major issues (e.g. impact transfer) before large scale

implementation; facilitate future planning, research, design and guide investment.

The above points are corroborated by the laboratory personnel. Multiple users highlighted that the need for environmental assessments in addition to the techno-economic analysis is primarily based on increasing requests by clients. These clients mainly are industries or organisations working at different levels such as governmental, intergovernmental, non-governmental, national or international. According to the discussions with the laboratory personnel, following are the reasons why environmental assessment is requested.

First is the quantification of environmental impacts of the use of new technologies. This enables comparison with the conventional ones. Also it gives an estimation and understanding of their environmental impact. *"The clients do not want their investments to become a liability"* as User2 puts it. Since the technologies are still in pilot/demonstration scale, future concerns could be curbed or entirely avoided by directing R&D towards the significant issues. Second, it enables the communication of information regarding environmental impacts of products. User2 also notes an increasing trend in enterprises wanting to disseminate the environmental impacts to the consumers or other stakeholders.

Finally, environmental indicators provide another perspective in decision-making. It was emphasized by multiple users that, technical and economic indicators have a higher importance. However, environmental information can especially help when techno-economic indicators are inconclusive. As user User6 says: *"Mobility is difficult (to decide) right now since there are various options, hydrogen, electric, synthetic natural gas...this could be a problem decided by environmental impacts"*

1.2.2 Relevance of life cycle assessment (LCA)

Life cycle assessment is a widely used to quantify environmental impacts of energy technologies, conventional or renewable [Turconi2013]. It is a comprehensive analysis methodology taking into account resource use and emissions throughout the entire lifetime. It is the relevant tool in this context and it is already used in the laboratory for assessing the energy systems. However, the comprehensiveness of life cycle assessment also contributes to complexities in its usage, especially in the current context.

Taking the ILCD handbook as the reference, LCA is briefly described here [European Commission2010]. It consists of four main steps as seen in figure 1.8(a). First step is the definition of goal and scope, where the overall objective of the study, coverage of the investigation, limitations and target audience are declared as precisely as possible. Second step, is the modelling and data collection step, called inventory analysis. Arguably the most time and resource intensive step, where all data has to be collected and modelled. The data is recommended to be representative of the actual situation, technologically, geographically and temporally. While being consistent with the requirements defined in the goal and scope. At this step, the activities in the system boundary are translated to exchanges with the environment (biosphere flows). These flows are essentially resource use and emission produced by all the activities. A conceptual representation is illustrated in figure 1.8(b).

Third step, is impact assessment, where biosphere flows are translated into a particular negative (or positive) impact on humans or the environment. There exists multiple impact indicators targeting the areas of protection. Example can be, carbon footprint for climate change or eutrophication for water bodies. Similarly, there are also multiple impact assessment methods since there are multiple ways of modelling environmental stresses [Weidema2015].

Finally in the last step, interpretation, the model and the results are confronted with the defined goal in the first step. The scope and inventory are iteratively refined to meet the study goals. At the end of the iterative process, the interpretation step facilitates the derivation of robust conclusions or even recommendations based on the overall quality of data and models.

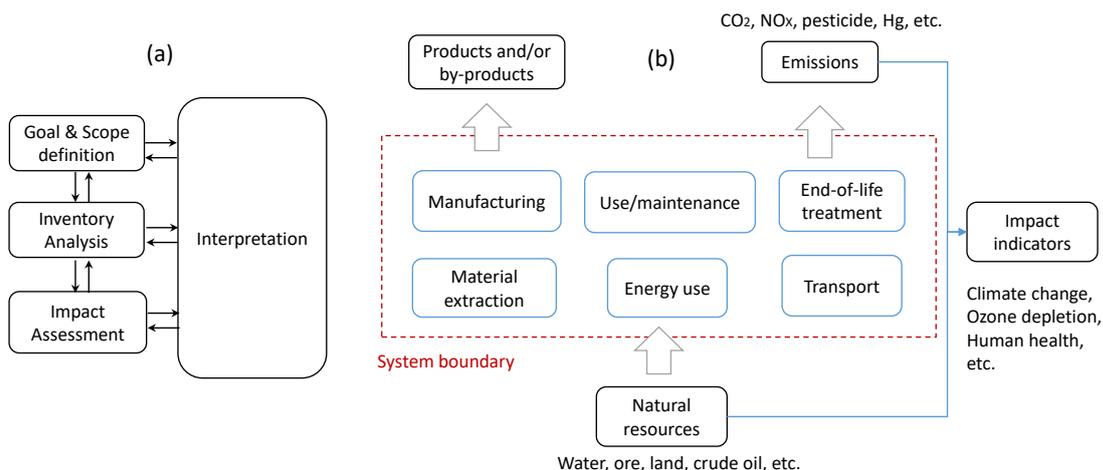


Figure 1.8: (a) Life cycle assessment framework adapted from ILCD handbook [European Commission 2010]; (b) A conceptual representation of inventory building and its use for impact calculation step of an LCA.

1.2.3 Laboratory engineers and life cycle assessment

As mentioned previously, LCA is occasionally conducted in the laboratory. In this subsection, information collected about the engineers regarding life cycle assessment is presented. Notably, their experience with LCA and preferences for development of an environmental assessment approach. This is especially useful since they would be the ones establishing the LCA models and conducting analysis on them.

The LCA competence amongst the lab personnel was found to be low. Majority of the participants were aware of the existence of the LCA methodology but had not conducted a study themselves. During the interviews it was found that only one user had experience with LCA investigations. The remaining 12 out of 13 users, only calculated the operating phase carbon dioxide emissions as a proxy for environmental impact. Apart from the carbon footprint, following indicators were mentioned: energy use, water use and resource use.

The personnel were also asked to quantify their preferences concerning the addition of environmental assessment. They were asked to rate importance they would assign to conducting an environmental assessment out of 10, with 10 being very important. Results are presented in figure 1.9. Consistency of models and reliability of results for a wide range of systems has the highest average rating. Followed by ease of data collection and use respectively. Regarding the ease of use, it was iterated by multiple users that they are willing to undergo additional training if necessary. Parametric and analysis choices available to the users allowing them to see effects of various parameters, trailed in just behind. It means enabling users to have ample choices to see results of their design. Finally, the time aspect of conducting an environmental assessment, is rated lowest since it depends on the particular project. It is related to the accuracy, since more the time spent, accuracy of the models should be higher.

1.3. Problems in the status-quo with respect to the inclusion of environmental impacts

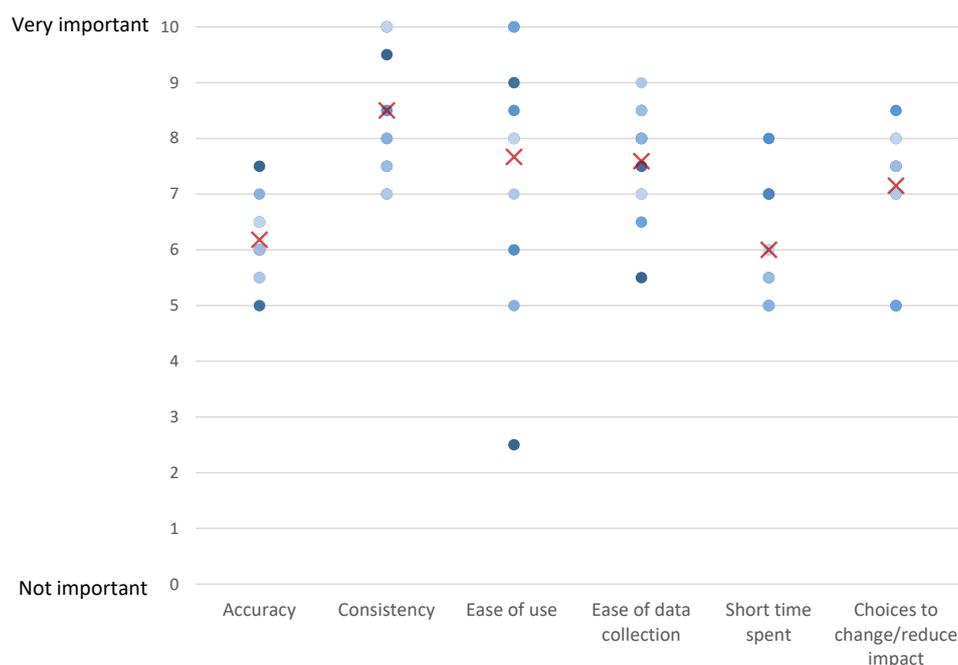


Figure 1.9: User preferences from the environmental modelling rated from 1-10 with 10 being very important. Blue dots represent the user rating and the red cross represents the average.

1.3 Problems in the status-quo with respect to the inclusion of environmental impacts

The primary responsibility of the laboratory is to design and plan different projects often involving emerging technologies using quantified technical, economic and environmental information. In doing so, the users in the lab aim to provide reliable analysis based on which decisions could be made on investments, R&D and even bring insights on policy-making.

With respect to the previous subsection, it is a clear that some members need LCA training. However, even if this limitation is improved, there are two problems found with the status-quo. It is argued that these two problems impede consideration of environmental information during the design and planning phase of hybrid energy systems:

1.3.1 Problem 1: Inadequate environmental impact estimation

First problem is insufficient environmental information for effective analysis. Environmental assessments for the projects in the laboratory are usually carried out in the following two ways.

Using emission values calculated either in Excel or added in the in-house simulations is the first approach. In order to do this, emission factors for processes are obtained from the ecoinvent database or literature. In this case, limited number of indicators (only global warming potential) and life cycle stages are included. In most of the cases, CO₂ emission calculations are limited to the operating phase. For emerging technologies, this approach risks impact transfer to other life cycle stages or to other impact indicators as concluded by [Laurent2012]. They demonstrated that shifting electricity production sources from fossil-based to renewables leads to a decrease in carbon footprint but the same trend does not appear for all other LCA indicators. Thus considering only carbon footprint during decision-making might lead to sub-optimal results. Especially when dealing with emerging technologies, consideration of other impact indicators to limit impact transfer is crucial.

Using specialised LCA software For the second case, when a more detailed study is required, LCA methodology as elaborated in section 1.2.2 is followed and a specialised LCA software, such as Simapro is used. Assistance is often requested from LCA specialists from outside of the lab. The main issue here is that the designers cannot see the consequences of changes in the design on the corresponding LCA indicators. The assessment is an iterative procedure, which involves multiple changes in parameters, input data, operation strategies and system architecture. Techno-economic information is readily available, thanks to the energy simulation software. On the other hand, since the environmental analysis is performed separately, environmental indicators are not treated in the same manner. Environmental analysis is thus done when the design has already been formalised. At this stage, LCA can only have limited inputs in design and planning.

For example, consider the case study presented in section 1.1.5. The influence of the efficiency of electrolyser on the environmental impact cannot be directly tested nor quantified using the current approach. Typically sensitivity analysis (local or global) is utilised for this purpose. Influence of parameters can be quantified clearly on the techno-economic indicators. For LCA indicators, using the current tools, only a local sensitivity analysis is possible, that too after a manual simulation of a system and then updating the LCA activities. Many such design parameters could be present in a system, insights on their behaviour in terms on environmental impacts are thus not readily available. Hence it is not possible to the users to provide appropriate analysis or recommendations on the design parameters and their influence on the environmental impacts.

Hence to summarize, either information is inadequate during the assessment (for the first case) or it is not utilisable with the relevant assessment tools (for the second case).

1.3.2 Problem 2: Exclusion of LCA indicators in the optimisation process

The second problem is that the LCA indicators have to be included as optimisation objectives if solutions with higher environmental efficiency are desired. This might be obvious to some readers but not so much for the others. Even if problem in the previous step is solved and appropriate LCA models are added, it is not guaranteed that it makes a difference to the solutions proposed by the optimisation algorithm. Here higher environmental efficiency is used instead of lower environmental impacts and is continued in the rest of thesis. Since it was observed that on a macro-level new energy technologies might have higher impacts than estimated due to increasing energy demand and/or rebound effect [Deng2017].

Here, a demonstration is made using the case study presented earlier to illustrate that LCA indicators have to be included in optimisation objectives to find solutions with higher environmental efficiency. The goal of dimensioning the system for hydrogen production remains the same but two components are added: solar PV and wind. The system configurations are searched with respect to the same techno-economic indicators (unmet load and levelized hydrogen cost). Thus, in theory there is potential to find system configurations with these two components and reduce the reliance on grid electricity. Consequently, it means that there is also a potential to reduce environmental impact per kg H₂ produced. Case study representation in the software can be seen in figure 1.10.

The meteorological data is obtained from renewable.ninja [Staffell2016,Pfenninger2016]. Linear cost model is used for the new components. Since LCA indicators are not taken into account during the search process itself, their estimation can be made at the end when the final solutions are obtained. This is also practical since LCA models are not yet developed, thus a rough estimation using only climate change potential is made at this stage.

The search is made with respect to the size range outlined in table 1.3. The genetic algorithm is allowed to run for 150 generations. Since there are 2 objectives, 2-dimensional pareto front is again obtained consisting of 156 different system configurations in figure 1.11. It can be seen that this curve is not different from the curve obtained in figure 1.6, where solar PV and wind

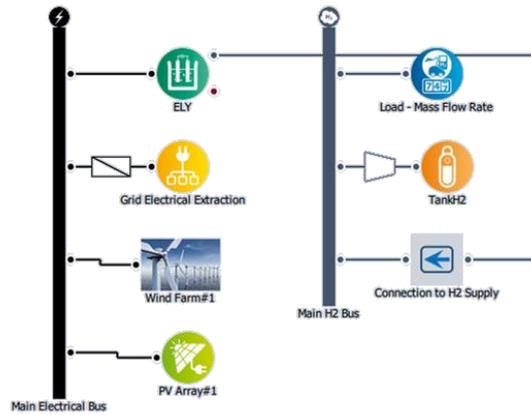


Figure 1.10: Representation of the case study in Odyssey

Table 1.3: Search range provided to the genetic algorithm.

	Min	Max
Electrolyser (MW)	0.5	2.5
Wind (MW)	0	8
Solar PV (MW)	0	8
H ₂ storage (m ³)	1	120

were not even included in the search process. On further analysis, it is found that there is no significant change in the proposed system configurations too. The maximum values wind and solar PV dimensioned in the pareto front solutions is 0 kW and 100 kW respectively. A summary of the system configurations obtained is presented in table 1.4. Since in this case study, grid electricity is comparatively cheap, configurations with wind and solar PV are a bit expensive and thus do not lie on the pareto curve. Even though solutions with them are desired, to have an option with different environmental impacts.

Table 1.4: Summary of the obtained 156 system configurations obtained at the end of the search

	Min	Max	Mean
Electrolyser (MW)	0.89	1.27	1.07
Wind (kW)	0	0	0.00
Solar PV (kW)	0	100	16.35
H ₂ storage (m ³)	12	43	26.18
Average grid use (kW)	920	1002	982.04

The climate change impact of hydrogen production from water electrolysis is primarily dependent on the type of electricity utilised as found by [Valente2017]. The climate change potential (kg CO₂-eq/kWh) of the three electricity sources: grid, wind and solar PV is 0.071, 0.014 and 0.07 respectively. These are generic values are obtained from ecoinvent 3.6 [Wernet2016] for France,

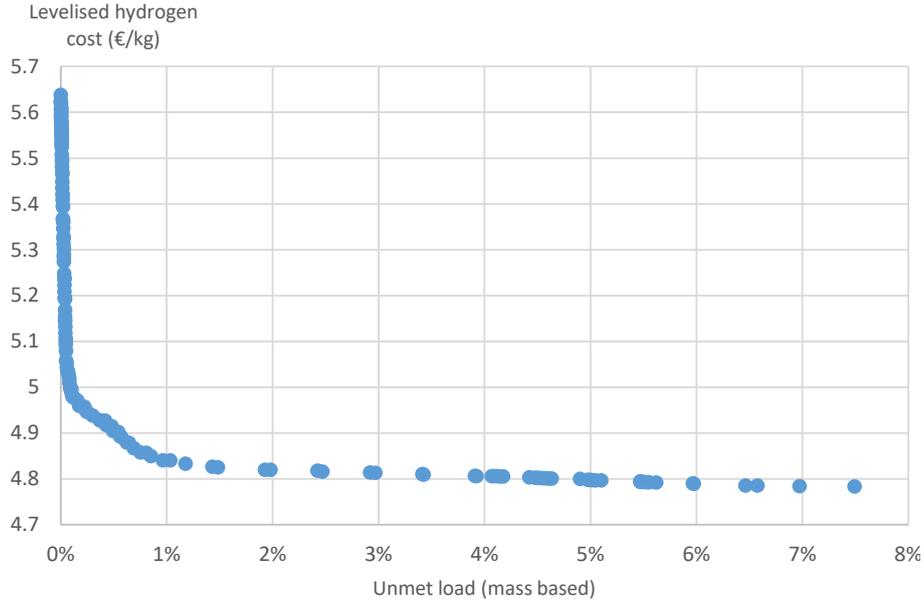


Figure 1.11: Pareto front of the results for the dimensioning problem

using the ILCD impact assessment method. It is clear from these values that when the electricity use from grid is minimised, carbon footprint of hydrogen production will also be lower.

For a demonstration, the impact of hydrogen produced based solely on the electricity impact can be calculated by the following equation:

$$I_{H_2} = \frac{LHV_{H_2}}{efficiency} * I_{electricity} \quad (1.1)$$

Where,

I_{H_2} : is the impact per kg hydrogen produced

LHV_{H_2} (33.3 kWh/kg H_2): Lower heating value of hydrogen

efficiency (56 %): of the electrolysis system

$I_{electricity}$: impact per kWh of electricity used

In the above equation, if climate change impact of grid electricity is substituted, it results in hydrogen production impact of 4.22 kg CO₂-eq/kg H₂. This rough estimation will be true for all of the system configurations obtained above since there is only a marginal inclusion of solar PV and furthermore, the difference with impact of grid electricity is negligible.

On the other hand if climate change impact of wind electricity is substituted in the equation 1.1, hydrogen production impact is reduced to 0.83 kg CO₂-eq/kg H₂. Thus we see more than 4 time difference when the electricity source is changed. This is of course a rough estimation of only one indicator and the final impact will depend on the system constraints, local weather conditions, end of life scenarios, etc. However, it can be seen that the search of system configurations enabling the reduction in carbon footprint of hydrogen is not possible in the present situation. Similar trend was noticed during multiple independent runs.

Here, if a reduction in the carbon emissions is desired, it is clear that more wind has to be included in the system. In the absence of LCA indicators included in the optimisation, wind can be manually sized in the system based on some estimations. Even then, it is quite difficult to

know if the performance of the system can be further improved. That is, it is probable that the points would not lie on or close to the real pareto front. This was also argued earlier when the relevance of multi-criteria optimisation was discussed. Thus it can be verified that systems with higher environmental efficiency cannot be reliably found unless environmental indicators are included as objectives in the multi-criteria optimisation.

1.4 Goals of the thesis

In the previous section it was concluded that there are two major problems preventing the inclusion of environmental impacts in the assessments of hybrid energy systems. Based on them, two main objectives of the thesis can be formulated:

1.4.1 Integrated LCA modelling along with hybrid energy simulations

In order to enable consideration of environmental impact, LCA models need to be integrated with the current techno-economic modelling in a combined approach. These models will ensure that instantaneous results are obtained based on a change in any of the system parameters. So as to allow users to effectively assess systems. An integration of LCA models in the existing assessment tools will also increase the probability that environmental impact is considered during investigations. Following points should be taken into account during the development of such models:

User profile - The users are experts in techno-economic analysis of hybrid energy systems but they are beginners in life cycle assessment. Developing LCA models in the context of emerging technologies needs high level of expertise and insights on the future evolution. It will therefore be more efficient to empower the users to develop the LCA models themselves rather than outsourcing the LCA analysis. Especially since technical evolution is closely linked to the changes in LCA results. It will leverage the already available competences in the laboratory. Thus it is important to keep the user profile in mind while building a general methodology of LCA models.

Prioritisation of data collection - Importance has to be assigned to the different data to be collected in the development of LCA models due to two main reasons. First is the variable level of detail required. According to the discussions with users, the granularity of LCA models depend on the resources available within the frame of different projects. The presently available techno-economic models also have a varying level of detail. The second reason is the presence of many uncertainties. LCA in itself is a data intensive process, with the 'future' scope expected in many investigations, amount of resources required increase exponentially. For example, background processes (e.g. recycle rate, electricity mix) have to be adapted to the future scope, lab/pilot scale inventory processes have to be adapted for a future industrial production, scaling equipment size from prototype to commercial scale, etc. If the importance of data to be collected is not quantified, due to limited time/resources available, practitioners will arbitrarily search for data. Consequently, disproportionate amount of resources might be allocated to modelling.

Compatibility with mathematical tools - As mentioned earlier, two mathematical tools are used often in the laboratory: global sensitivity analysis and genetic optimisation. To ensure that the environmental information is treated in a similar way as other performance indicators,

it is important that LCA models too are compatible with these tools. It will increase the chances that the new indicators are taken into account by the users.

1.4.2 Inclusion of LCA indicators in multi-objective optimisation

It was noted that, environmental impact adds an extra dimension in decision-making. Multi-criteria optimisation is already used to obtain solutions with trade-offs between the techno-economic objectives. Trade-offs with further environmental indicators will be useful for design and planning of systems while including the environmental impacts. Furthermore it will be practical to the users to have additional information more or less in the same structure as currently utilised.

It was demonstrated that without including LCA indicators as objectives in the multi-criteria optimisation, solutions with higher environmental efficiency cannot be searched. The problem then is, LCA can have up to 20 different objectives unlike one indicator each for tech-eco analysis. The maximum objectives that could be included in Odyssey presently is three. Hence a strategy to effectively include LCA results in optimisation is required.

1.5 Research Approach

Once the thesis goals were formulated, the approach followed, was to search the relevant state of the art topics to solve them. This led to the identification of topics in the literature that were missing or ones that could be improved further. Consequently, research proposals were concocted to address these issues. The adopted research approach is described further below, and a simplified illustration of the approach can be seen in figure 1.12.

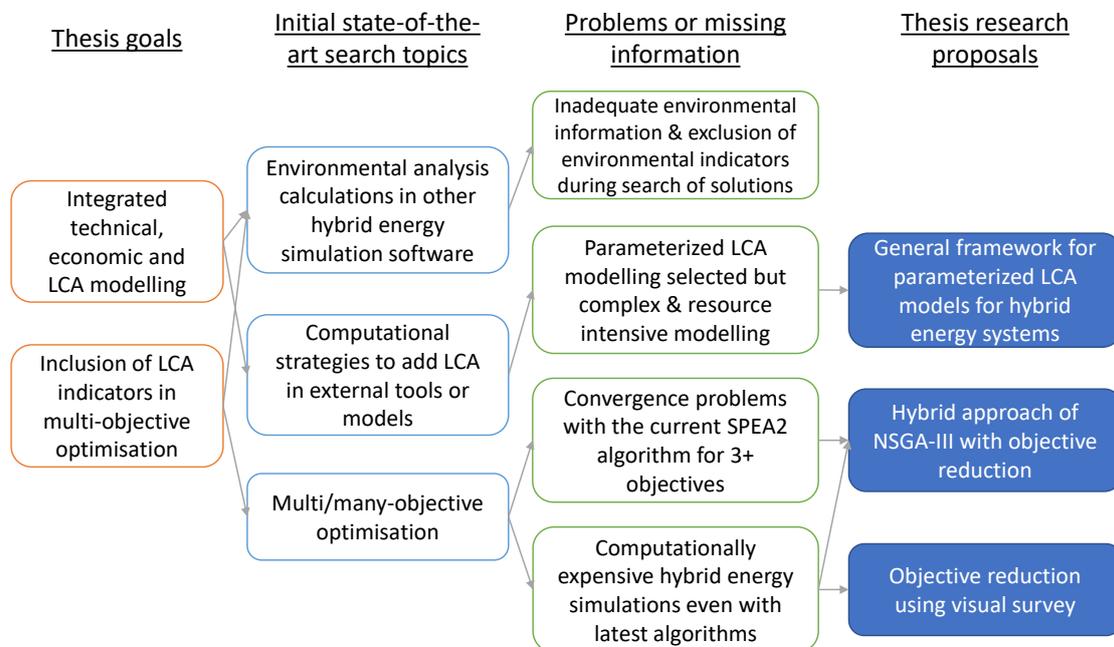


Figure 1.12: A simplified illustration of the research approach followed in this thesis

As we saw in this chapter, two key problems were identified after discussions with laboratory personnel and utilisation of tools in the lab. These two problems prevent consideration of environmental information in design and planning of hybrid energy systems. The two primary goals of the

thesis could then be formulated: integrated modelling including LCA models and multi-objective optimisation including the LCA indicators.

In order to solve these goals, state-of-the-art topics were searched. First step was to see how environmental information is calculated and included in other hybrid energy simulation software. It was found that both of the two problems also exist in them.

The next step was to search for computational strategies utilised in the literature in order to integrate LCA with other tools or software. Since it would help get an overview of the possible approaches for integrated modelling with LCA. From this exercise, it was concluded that parameterized LCA modelling could be a valid approach to integrate the LCA models with the energy simulations. However, there were two problems that remained unsolved in the context of hybrid energy simulations. First, the modelling is complex since the process requires high level of expertise in both topics: energy systems and LCA. Second, due to emerging technologies often appearing in this context, the modelling is resource intensive with many uncertainties.

Thus to solve these issues, *a general framework for establishing parameterized LCA models for hybrid energy systems* is proposed. For this purpose, an initial methodology was built by anticipating the issues that could be faced while developing parameterized LCA models for emerging technologies. This initial methodology was then given to an energy engineer to implement a preliminary model of the PEM electrolyser. Based on the feedback during this process, the methodology and the LCA model were improved. Other elements of the literature such as global sensitivity analysis and emerging technologies were also consulted during this process.

Then to solve the second problem of including LCA indicators in the search of solutions, the literature on multi/many-objective optimisation was consulted. It was found that the currently utilised algorithm SPEA2 might have convergence problems when optimisation with respect to 3+ indicators are desired. Given the experience and the past research laboratory personnel, a new generation evolutionary algorithm NSGA-III which does not have these issues was chosen.

During the literature search, objective reduction techniques were also found. These techniques basically identify redundant optimisation objectives to be temporarily discarded, which do not change the pareto front but increase the efficiency of optimisation algorithms. These are quite relevant since LCA indicators were found to be correlated in various investigations in the literature. Moreover since the hybrid energy simulations are computationally expensive, they could be quite helpful in reducing the computational time. Thus a *hybrid approach of NSGA-III with objective reduction* is proposed.

Finally, during the implementation of the above approach, it was found that LCA indicator correlations could be found before the optimisation process based on average estimations. Given that the system satisfies certain conditions. This finding was later finalised as the proposal of, '*Objective reduction in LCA using a simple visual survey*'. The rest of thesis unfolds in the following manner:

Chapter 2 elaborates the findings from the state-of-the-art. The above discussed topics are justified as much as possible followed by the formulation of research proposals of this thesis.

Chapter 3 then describes the proposal of establishing parameterized LCA models. The methodology is demonstrated by applying it on a PEM electrolyser case study.

Chapter 4 outlines the proposal of the proposed objective reduction technique using a bike design case study. The results are verified using principal component analysis, a well established objective reduction technique in the literature used for the same purpose. Recommendations are provided on its utilisation.

Chapter 5 presents the final proposal of the hybrid NSGA-III approach with objective reduction. Performance of the new approach is tested with the original algorithm using a dimensioning case study of hybrid energy systems. Recommendations are then made on the use of this approach.

2

Positioning the thesis in the existing literature

Summary: In this chapter, the state-of-the-art topics are reviewed in order to solve the formulated goals of the thesis. First activity was to investigate the available hybrid energy simulation software. The main objective was to assess how the environmental information is calculated. Then, in the second section the LCA integration strategies in the non-LCA software or models are investigated. So as to formalise the type of LCA modelling to be used in the thesis. Literature on the inclusion of LCA indicators in the optimisation process is consulted in the following section. This enables the definition of the optimisation strategy to be used in order to include the LCA indicators in the search of solutions. Finally, from the disparity found between the elements required to satisfy thesis goals and the literature, research proposals are formulated in the last section.

Contents

2.1 Environmental information in the hybrid energy simulations	25
2.1.1 Summary: Insufficient environmental information	26
2.2 LCA computational models	27
2.2.1 Computational structure of LCA	27
2.2.2 Search Method	28
2.2.3 Selection criteria	29
2.2.4 Results	29
2.2.5 Summary: Parameterized LCA models are selected but some issues remain unsolved	37
2.3 Inclusion of LCA indicators in optimisation	38
2.3.1 Problems with the currently used genetic algorithm SPEA2	38
2.3.2 Use of genetic algorithm NSGA-III in this thesis	39
2.3.3 Optimisation overview in life cycle assessment	43
2.3.4 Objective reduction techniques in life cycle assessment	44
2.3.5 Summary: A possible combination of NSGA-III with objective reduction?	47
2.4 Synthesis of research proposals	48

2.1 Environmental information in the hybrid energy simulations

One of the first activities of this thesis was to compare the environmental analysis in the hybrid energy simulation software [Sharma2019]. The objective of this exercise was to investigate whether the two requirements formulated in the previous chapter were already satisfied in the other similar software. A comprehensive list of more than 50 software was compiled using energyplan.eu and investigations by [Sinha2014, Connolly2010]. These software were screened using so that the relevant ones are retained using the following criteria:

- Software catering to a geographical area greater than a region were excluded.
- Only the software dealing with technical or economic simulation were selected. Hence, the ones dealing with objectives such as policy, markets, networks, etc. were neglected.
- Finally, software unavailable on the market because of being obsolete or internally used, were not investigated in this paper.

Five software passing the screening are HOMER ¹, iHOGA ², EnergyPRO ³, RETScreen Expert ⁴ and TRNSYS ⁵. Amongst them, Odyssey can be added as well.

As we see in table 2.1, only one software includes environmental information of more than one life cycle stage. Furthermore, only limited emissions are taken into account. This is inappropriate for systems involving renewables or energy storage technologies, where most of the impacts are from upstream or downstream stages.

Table 2.1: Summary of the environmental analysis in the reviewed software [Sharma2019].

Software	Life cycle stages considered		Emission factor units	Emissions						Possibility to add more emissions	Emission limits or penalties
	Operation	Other		CO ₂	NO _x	SO ₂	CO	Unburnt fuel	Particulate matter		
HOMER	✓	-	kg pollutant per unit fuel consumed	✓	✓	✓	✓	✓	✓	-	✓
iHOGA	✓	✓	kg pollutant per unit fuel consumed OR kg pollutant per unit size of equipment	✓	-	-	-	-	-	-	-
EnergyPRO	✓	-	kg pollutant per unit fuel consumed	✓	✓	✓	-	-	-	✓	✓
RETScreen Expert	✓	-	kg pollutant per unit energy produced	✓	✓	-	-	✓	-	-	-
TRNSYS	-	-	-	-	-	-	-	-	-	✓	-
Odyssey	✓	-	kg pollutant per unit fuel consumed	✓	-	-	-	-	-	✓	-

A demonstration is seen in figure 2.1, where a comparison between electricity production from wind turbines, solar PV, conventional natural gas and coal plants according to their climate change potential. The resulting impacts are distributed along the life cycle stages. For solar PV and wind turbines, data was assembled from a paper investigating 153 life cycle assessment (LCA) studies [Nugent2014]. While for natural gas and coal, it was taken from LCA review and harmonisation studies [O'Donoghue2014, Whitaker2012]. It is clear that, for conventional technologies, most of

¹ <http://homerenergy.com/>

² <http://ihoga-software.com/>

³ <https://www.emd.dk/energypro/>

⁴ <http://www.nrcan.gc.ca/energy/software-tools/7465>

⁵ <https://sel.me.wisc.edu/trnsys/>

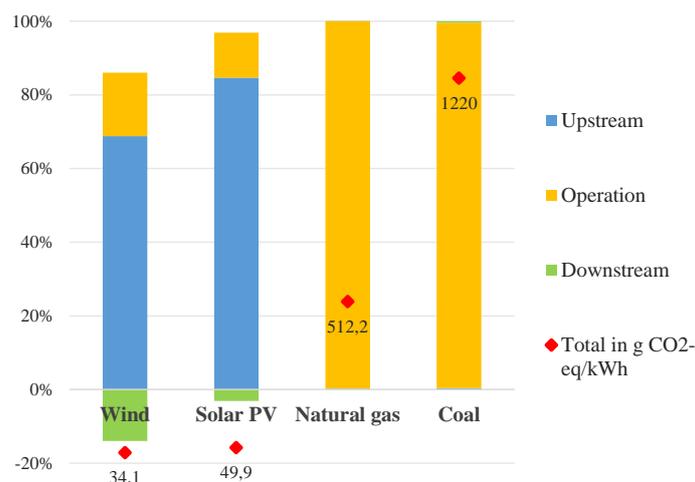


Figure 2.1: Breakdown of climate change potential according to different life cycle stages for electricity production from wind turbines, solar PV, conventional natural gas and coal fired plants [Sharma2019].

the climate change impact results from the use phase. These two studies also show that first order estimates of global warming potential is possible only with a knowledge of certain supply chain steps of the fuel cycle. However, the same is not true for renewable technologies, where most of the impacts come from the manufacturing/end-of-life phases. The latter have a lower carbon footprint by an order of magnitude, than conventional technologies but this might not be true for other indicators such as mineral resource depletion. Hence, decisions based only on greenhouse gas emissions might lead to sub-optimal conclusions due to impact transfer [Laurent2012].

Coming to the functionalities offered by these software, 4 offer a possibility of searching system configurations according to desired indicators. Out of them, EnergyPRO is used for planning the operation of the system. For example, when to operate a component according to the weather, electricity market, etc. It does not offer the possibility of sizing the system. HOMER does offer searching solutions with respect to a single criteria. It does this by simulating all possible combinations followed by sorting the solutions in an increasing or decreasing order. A summary can be seen in table 2.2.

Only 2 software (iHOGA and Odyssey) offer search of solutions with respect to more than one objective. Since they use genetic algorithm, their convergence is likely to be faster than HOMER even for single objective optimisation. In terms of including environmental indicators in the search, only CO₂ emissions are included in them that too for limited components and life cycle stages.

2.1.1 Summary: Insufficient environmental information

It is concluded that environmental information provided in them is insufficient to support decision-making in the context of this thesis. Main reason being that the analysis is not sufficiently adapted towards the characteristics of new energy technologies. Impact indicators and life cycle stages are excluded from estimation. Thus the two outlined problems of insufficient impact estimation and exclusion of environmental indicators also remain in the other similar software.

Table 2.2: Summary of the environmental analysis in the reviewed software [Sharma2019].

Software	Simulation		Sizing	Operation strategy optimization	Computation type
	Economic	Technical			
HOMER	✓	✓	✓	✓	Simulation + mono-objective search
iHOGA	✓	✓	✓	✓	Simulation + Multi-objective genetic algorithm
EnergyPRO	✓	✓	-	✓	Simulation
RETScreen Expert	✓	✓	-	-	Simulation
TRNSYS	-	✓	-	-	Simulation
Odyssey	✓	✓	✓	✓	Simulation + Multi-objective genetic algorithm

2.2 LCA computational models

The goal of this section is to identify LCA modelling strategies in the literature and the ones that could be undertaken in order to solve the first problem. Thus the objective of this section is to identify strategies in the literature that could be used to integrate LCA methodology with external non-LCA tools and models. More specifically in the context of the thesis, LCA modelling that could take inputs from energy simulation and calculate impact indicators. Thus, it is clear that quantitative results are required.

Apart from the conventional LCA computation, other strategies have also been used in the literature. Once these approaches are classified and characterised, an overview of the thesis goals with respect to already developed strategies can be obtained.

Before proceeding with the search, it is necessary to first specify the conventional computational structure of LCA. Hence, this is described next, followed by the literature search methods. Then from the found studies, reasons for retaining certain studies are elaborated. Finally, the identified strategies are classified and described.

2.2.1 Computational structure of LCA

In this section, the standard methodology to calculate inventory and impact assessment is elaborated as per [Heijungs2002]. Let's start with a technosphere matrix \mathbf{A} , which consists of relationships between the economic activities with each other. Final demand vector \mathbf{f} is the reference flow of the LCA study. Where, the amount of each activity desired for the LCA calculations are specified. In order to satisfy the demand \mathbf{f} , scaling vector \mathbf{s} for each of the activity is the technosphere matrix is defined. Since \mathbf{A} (often, a LCA database) and \mathbf{f} (demand) are known, scaling factors can be calculated in equation 2.1 using matrix operations.

$$\mathbf{A} \cdot \mathbf{s} = \mathbf{f} \iff \mathbf{s} = \mathbf{A}^{-1} \cdot \mathbf{f} \quad (2.1)$$

As a fictional example, two activities of electricity and fuel production are shown in the technosphere matrix. For 10 kWh electricity production, 2 litres of fuel is required. While for producing 100 litres of fuel, electricity is not required. Let's assume 5 kWh of electricity as the reference flow.

$$\begin{array}{l} \text{Electricity prod.} \quad \text{Fuel prod.} \\ \text{Litre of fuel} \\ \text{kWh electricity} \end{array} \begin{pmatrix} & -2 & 100 \\ & 10 & 0 \end{pmatrix} \times \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 5 \end{pmatrix}$$

Hence, the scaling factors would be 0.5 and 0.01. They are then used to calculate impact of each activity using the environmental intervention or the biosphere matrix **B**. It has the same number of columns as matrix **A**, meaning the same activities. The rows represent the 'exchanges' of each activity with the environment. The total environmental exchange **g** for the reference flow is then found using the below equation:

$$\mathbf{g} = \mathbf{B} \cdot \mathbf{s} \quad (2.2)$$

As an extension of the above example, lets say fuel production requires extraction 50 litres of crude oil and emits 20 kg each of CO₂ and SO₂. Whereas, electricity production does not require crude oil directly but emits 30 kg and 40 kg of CO₂ and SO₂ respectively. Hence, equation 2.2 can be represented as:

$$\mathbf{g} = \begin{array}{l} \text{Crude Oil} \\ \text{CO}_2 \\ \text{SO}_2 \end{array} \begin{array}{l} \text{Electricity prod.} \quad \text{Fuel prod.} \\ \begin{pmatrix} 0 & -50 \\ 30 & 20 \\ 40 & 20 \end{pmatrix} \end{array} \times \begin{pmatrix} 0.5 \\ 0.01 \end{pmatrix} = \begin{pmatrix} -0.5 \\ 15.2 \\ 20.2 \end{pmatrix}$$

Once all the environmental exchanges are calculated, they are multiplied by the characterisation matrix **C** to get the impact indicators, **I**.

$$\mathbf{I} = \mathbf{C} \cdot \mathbf{g} \quad (2.3)$$

Here, the rows of matrix **C** represent the selected impact assessment methods. While the column represents the characterisation factor for each environmental intervention. Thus the number of columns of **C** is equal to number of rows of **g**. Finally, *I*₁ and *I*₂ are the two impact indicators for the reference flow of 5 kWh electricity.

$$\mathbf{I} = \begin{array}{l} \text{Impact 1} \\ \text{Impact 2} \end{array} \begin{array}{l} \text{Crude oil} \quad \text{CO}_2 \quad \text{SO}_2 \\ \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \end{pmatrix} \end{array} \times \begin{array}{l} \text{Crude oil} \\ \text{CO}_2 \\ \text{SO}_2 \end{array} \begin{pmatrix} -0.5 \\ 15.2 \\ 20.2 \end{pmatrix} = \begin{pmatrix} I_1 \\ I_2 \end{pmatrix}$$

This is a simplified representation of the calculations, they are conducted at a much larger scale in LCA investigations. For example, no. of activities in ecoinvent 3.6 database are 13831, which translates to the 13831 rows and columns for the matrix **A**.

2.2.2 Search Method

A preliminary search was carried out using the keywords 'Simplified LCA' in google scholar and scopus. During this search, LCA models were found, combined with external 'non-LCA' tools, models or databases. Such LCA investigations are often termed as 'simplified or streamlined LCA study'. The objectives of these investigations was generally to provide decision support or for design. Similarly, simplification techniques conventionally utilised by LCA practitioners were also found.

Consequently, to focus on studies in-line with the objective, the search string comprising of keywords for: LCA, simplifications and integration with other models was formulated. The search string presented in figure 2.2, resulted in 920 results in Scopus. These results consisted of books, journal and conference publications.

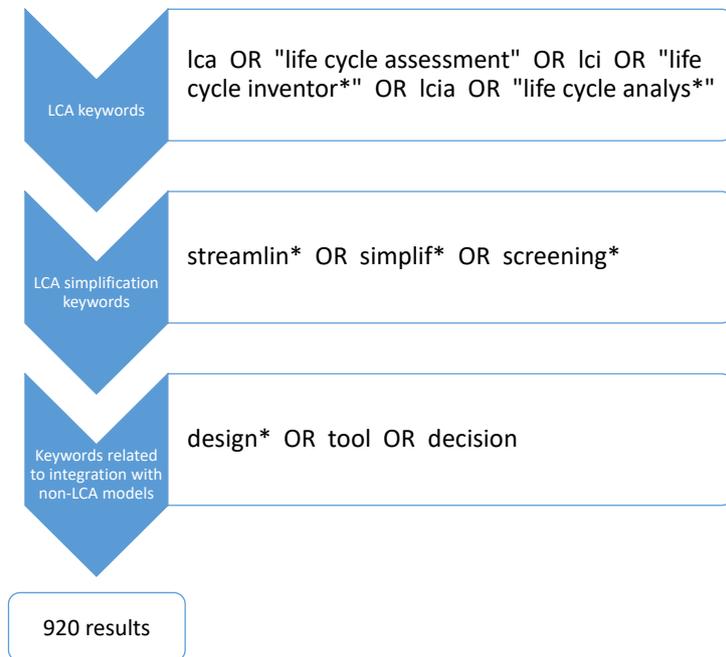


Figure 2.2: The search string used to identify documents in Scopus.

2.2.3 Selection criteria

Abstracts and/or full text of these 920 results were then reviewed. Thirty eight studies could not be pursued further since their full-text was not available online or they were not written in English. Further, 791 were discarded since they did not fit in with the objective of the search due to one of the following reasons:

- They were not an environmental LCA study
- They did not integrate LCA with an external tool or did not present an approach which could be utilised to simplify LCA computations
- They calculated qualitative LCA impact indicators

As a result of the selection, 75 studies were retained and a further 91 studies were added through snowballing, resulting in a total of 165 studies classified.

2.2.4 Results

Statistics regarding the total 165 studies analysed in this subsection are presented in figure 2.3. An increasing year wise trend can be seen in these kind of investigations. In terms of the field of application, built-environment or construction has the maximum applications. Often studies involving an integration of building databases with LCA database were noticed. In the thesis of [Tasala Gradin2020], the lack of data was identified as the main motivating factor for simplification. They review all the simplification strategies in LCA. However, for the studies found here, the main objective was quick estimation of impact for design or decision-making. LCA was often

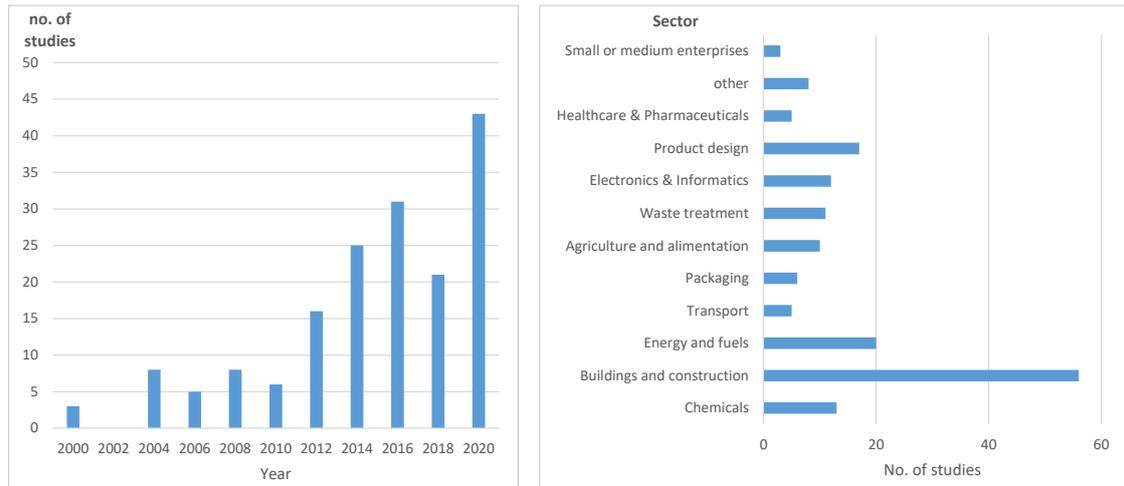


Figure 2.3: Statistics of the total 165 references analysed. (left) Year-wise distribution; (right) Sector-wise distribution.

integrated with a non-LCA tool or even as a simplified excel sheet to guide stakeholders about the environmental impact.

Globally, three ways of computing LCA indicators were found combined with other tools/models: Parameterized modelling, Functional approximation of LCA calculations and Data mining techniques. The basic calculation principle behind each of them, examples of studies found under this classification and relevance of each approach is presented below.

Parameterized LCA models

Concept For parameterized LCA models, definition from [Heijungs2020] is preferred, as it best defines the studies segregated under this category. They are the ones where LCA process data (or a part of it) depend on one or more settings. The settings are often the relevant parameters $(p_1, p_2, p_3, \dots, p_d)$, using which the demand vector $\mathbf{f} = (f_1, f_2, \dots, f_l)$ is calculated. The output of the parameterized model would thus be the amounts of different activities required in the scope of the overall LCA model:

$$\mathbf{f} = \psi(p_1, p_2, p_3, \dots, p_d) \quad (2.4)$$

Once constructed, the parameterized models allow the calculation of impacts according the change in the respective parameters. The LCA calculations are then normally carried out through equations 2.1 - 2.2. These steps could be quite computationally expensive due to the large sizes of the matrices. Recall from section 2.2.1, the impact indicators can be calculated as:

$$\mathbf{I} = \mathbf{C} \cdot \mathbf{B} \cdot \mathbf{A}^{-1} \cdot \mathbf{f} \quad (2.5)$$

A part of the above equation could also be pre-calculated if the activities present in \mathbf{f} are known. These are then known as precalculated aggregated datasets [Lesage2018]. Impact per unit of each activity in \mathbf{f} is calculated separately. Then, the vector of 'n' impact indicators $\mathbf{I} = (I_1, I_2, \dots, I_n)$ can be calculated as the sum of impacts of each activity in the system:

$$I_n = \sum_l I_n^l * f_l \quad (2.6)$$

Where, I_n^l is the impact per unit of activity 'l' for an impact category 'n' which is pre-calculated. Similarly, a parameterized model could also be defined in terms of the environmental exchanges in matrix \mathbf{g} :

$$g_h = \sum_l g_h^l * f_l \quad (2.7)$$

Where, g_h^l is the pre-calculated environmental exchange 'h' per unit of activity 'l'. Impact indicators are then calculated using equation 2.2. This is especially useful when a limited number of activities are used each time. The calculation of large matrices is thus avoided, making the process more efficient. Another important point is that the pre-calculated datasets also enable the mathematical formulation of LCA models which could be useful various in contexts, such as analytical sensitivity analysis methods or multi-integer linear programming.

For instance, consider the electricity production from the example defined in section 2.2.1. If let's say climate change impact (I_{CC}^{elec}) of the electricity is precalculated to be equal to 0.5 kg CO₂ eq/kWh, this value can be directly multiplied by the electricity consumed which is presumably calculated by a parameterized model.

Application in LCA No. of studies classified under this category was 131. Similar to any LCA study, the level of granularity of parameterized LCA vary in the literature. This type of integration has been done in [Morbidoni2010] who combined product design tools with the LCA specialised software OpenLCA. It was often found that the demand vector in equation 2.4 was calculated by an external model. Then its entry to an LCA software/calculations was automated. The open source LCA tools such as OpenLCA [Ciroth2014] and Brightway2 [Mutel2017] have further facilitated this step.

LCA calculations can also be included with other analysis tools. For example, the BIMEELCA tool is developed for an automatic LCA during building design [Santos2019, Santos2020]. Different levels of detail of an LCA are defined depending on the life cycle stages included in the scope. Pre-calculated aggregated datasets are used for impact estimation in this case.

Use of area-specific databases or using limited number of activities from a database are also common in these cases. For example, use of Okobaudat [Gantner2018] in built-environment LCA tools [Zimmermann2019]. Aggregated datasets have also been used with optimisation algorithms to select materials [Ng2020] or to assess investment in energy sources [Tietze2020].

This type of formulation has also been widely used to make area-specific decision or assessment tools for non-experts in LCA. Often only a handful of inputs are required by the tool for a high-level estimation of impact. The objective of these tools is often to provide quick environmental information to the relevant actors. These tools have been developed for instance in solar heating and cooling systems [Beccali2016], milk production [Pirlo2019], urban water industry [Schulz2012] and packaging [Verghese2010].

Relevance Since the basic computational structure of these models remains the same as in the conventional LCA, similar accuracies can be expected. Thus, the relevance of the results will then depend on the scope of the models and the data quality. The simplifications utilised in the conventional LCA studies are also applied in this case. Some examples are: use only of databases, narrowing the range of environmental impacts considered, neglection of upstream/downstream processes, etc. Hence, conclusions that could be derived from these models, depends on their modelling assumptions.

Functional approximation of LCA calculations

These models basically estimate the outputs (LCA indicators) with respect to the different inputs (parameters) like the models in the previous sections. However, the conventional LCA calculations

as seen previously are by-passed completely. The estimations are made by establishing mathematical relationships between the inputs and outputs leveraging previously conducted LCA studies. In theory, any mathematical approach predicting outputs with respect to the inputs of a could thus be utilised. The calculations under this category could thus be represented by the following equation:

$$\mathbf{I} = \phi(p_1, p_2, p_3 \dots p_d) \quad (2.8)$$

In the LCA literature, two techniques were found so far: multi-linear regression and artificial neural networks.

Multiple linear regression

Concept It enables the modelling of relationship between the response and one or more independent variables [Iguar2017]. A model predicting environmental impact I_n with respect to m independent variables can be defined in the following equation:

$$I_n = a_0^n + a_1^n p_1 + a_2^n p_2 \dots a_m^n p_d \quad (2.9)$$

The predictor variables can be either selected using literature review, industrial experience or statistical methods. LCA indicator data has to be made available with respect to each of the selected variables/parameters p_m^n . The data can be empirical or artificially generated. The most common method to calculate the co-efficients (a_m^n) is the ordinary least squares estimator. This can be easily done in Excel or in programming languages such as Python or R. Later, the impacts of new cases are estimated by entering the predictor variables. As we can see from the above equation, 'n' separate equations need to be formulated for 'n' impact indicators.

Application in LCA No. of studies classified under this category was 22. Roches *et al.* use this approach to extrapolate impact of agricultural products in different countries using nine farming inputs such as fertiliser use, machine use, etc [Roches2010]. Emission factors (CO₂/kWh) of electricity production from 444 worldwide coal plants were estimated by [Steinmann2014]. Along with multiple linear regression, they also used *local* linear regression for model fitting. They hypothesized that 5 predictor variables explain the variability in the emission factors: plant age, capacity, coal type, steam pressure and GDP per capita of plant location.

Carbon footprint of wafer fabrication was calculated using 3 parameters selected through literature review, industrial expertise and process & correlation analyses [Huang2016]. Padey *et al.* first developed a comprehensive parameterized model of wind turbines. Then, they applied GSA to identify the two most influential parameters as lifetime and load factor on the global warming impacts. A simple regression model for quick CO₂ emissions estimation from wind electricity based on these two variables was then developed [Padey2013].

Lullo *et al.* also used GSAs to effectively select the important parameters for their regression models. They further make available an Excel workbook to convert excel-based LCAs into regression models [Di Lullo2020]. Regression was also used to predict indicators but based on other known indicators by [Pascual-González2015]. This approach was applied to oil and electricity product categories of Ecoinvent database. They were able to achieve accuracies of more than 80% using 5 indicators as the learning set for electricity and more than 85% using 3 indicators for oil datasets.

Relevance The main advantage of such models is the ease of utilisation. They could be used simply as an equation, that can account for different variabilities in the impact. Hence, avoiding the LCA calculations. Furthermore, they could be used where LCA needs a mathematical formulation.

In terms of the relevance of results, it highly depends on the scope of the model and the values to be predicted. For example, [Padey2013] reported an accuracy of more than 85% for global

warming impact from wind electricity. On the other hand [Roches2010], report relatively high accuracy (83%) for some indicators less than 67% for others. Hence, it has to be clearly defined under what conditions the model will offer relevant results. Next, the 'goodness' of the fit measures such as co-efficient of determination (R^2) can be used to evaluate the representativeness of the data. Thus in general, they could offer initial estimates or data for 'screening LCA' depending on certain conditions such as goodness of the fit, type of data point, etc.

Artificial neural networks (ANN)

Concept They are being widely used for recognition of image, video, language and sound [Wolfgang2011]. Furthermore, ANNs can also display basic creativity in generating these patterns. Their relevant application in this context is function approximation, i.e. the non-linear predictions after *learning* the relationships between input parameters and outputs. In the context of LCA, they could be defined as equation 2.8 but not further. They are the black-box models predicting the impact with respect to some parameters or settings.

Similar to regression, output data with respective variation of input variables is necessary for neural network models. An ANN is made up of multiple layers of inter-connected artificial neurons, illustrated in figure 2.4 on the left. Each neurons in a layer is connected each neuron in the adjacent layers. However, neurons belonging to the same layer are not inter-connected [Skansi2018]. A neuron receives two types of inputs for calculations: weighted values (see figure 2.4 on the right) from all neurons in the previous layer and an additional value of *bias* (b). The weights are not necessarily bound between 0-1. The output of the neuron could simply be a sum of these inputs for linear neurons. However, in most cases, these inputs are transformed with the help of a non-linear function. The most commonly used being a sigmoid or logistic function, illustrated figure 2.4 on the right. The calculated output is then multiplied with the respective weights and sent to all the neurons in the next layer [Skansi2018].

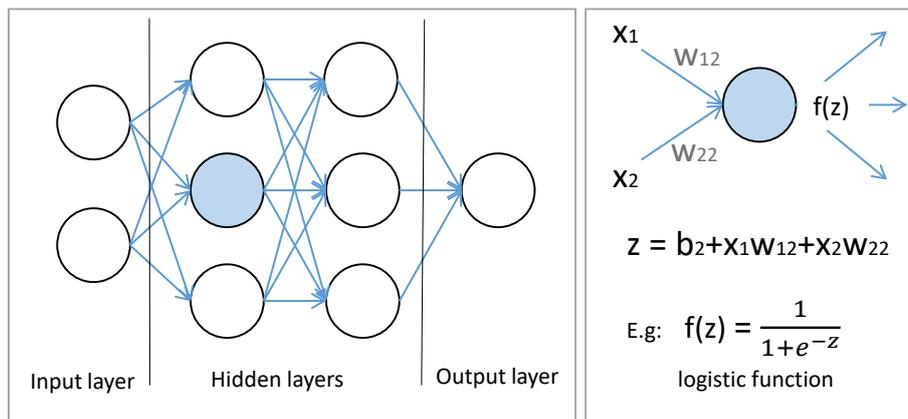


Figure 2.4: (Left) Simplified illustration of an artificial neural network; (Right) An artificial neuron, simplified., illustration adapted from [Skansi2018]

Initially, the network *learns* mapping of output from inputs using the training data. This process is essentially the network adjusting the weights (w) to ensure the error between the predicted and the actual value is low. The ANN might require multiple iterations through the training data to arrive at an acceptable error value [Kalogirou2001]. Once the error is below the acceptable values, weights of the network are held constant in order to predict outputs for testing, validation and eventually for predicting new data points. From a practical point of view, ANN implementation is

offered by many open source libraries available in programming languages such as Python, Java or C++.

Application in LCA No. of studies classified under this category was 8. Park & Seo used this approach to predict early-stage impact of household products using product attributes such as power consumption, lifetime, material distribution, etc. with accuracies more than 72% [Park2003]. Wernet *et al.* applied ANN's to predict impact of chemicals at an early stage using their molecular structures [Wernet2008]. They further compared this approach in terms of accuracy and time saved in comparison with 3 other common approximation techniques used by LCA practitioners: extrapolation, use of proxy data and process models [Wernet2012]. A number of studies have thereafter developed this investigation further taking into account other factors such as thermodynamic properties [Song2017, Karka2019, Kleinekorte2019]. Khanali *et al.* trained ANN's by collecting on-field data of Tea production in Iran followed by analysis of predictions to get accuracies of more than 85% [Khanali2017].

Relevance ANNs are especially useful when the data is detailed, complex, non-linear and highly co-related. In contrast to the regression models, where the exact mathematical relationship has to be specified for curve-fitting (e.g. order of polynomial), ANNs calculate the relationships by themselves. As a downside, they are even more opaque to introspection than regression models. It is not possible to comprehend inner workings of ANN to find specific impact sources or contributions from a prediction. Furthermore, these models might require more resources (data, time and computational power) for their development as compared to other approaches. Their applications are also limited to the training dataset. Further the departure from it, lower is the accuracy.

Therefore the accuracy of ANNs depend on the similarities of the point to be predicted with the training set. A valid use of impact estimations based on ANNs could be filling data gaps. The instances where data of low importance is missing. By low importance, it is meant where the data is expected to have a low contribution to the impact. Moreover since the ANN models are opaque to analysis, estimation for a data point cannot be verified. Thus they are only recommended to be used to estimate the order of magnitude of the impact when the prediction point is similar to the training set.

Data mining techniques

Concept In brief, the techniques described below depend on calculating the impact of a new product based on its similarities with other products already present in the database. The similarities could be in terms of functions, structure, materials, behaviours or environmental impact. These similarities are determined using text descriptions in bill of materials, LCA databases or in a manually created database. Data mining and machine learning techniques are leveraged for this purpose. These studies often also involve facilitating the mapping of LCA databases with others, for example, a BOM database. Multiple data mining strategies have been cited in the 4 publications found under this category. However, in this section, attention is focused on two techniques, which form the basis of all the calculations in the found references: clustering and classification. Then, the studies are presented along with their LCA estimation method.

Clustering is defined as '*grouping together similar objects*' while separating dissimilar objects in separate groups [Bramer2007]. For example, grouping developing and developed economies. A salient point in this aspect is that the clusters are *unlabelled* [Skansi2018]. A visual representation of clustering for two dimensional data points can be seen in figure 2.5. The algorithms find clusters typically when distance between neighbouring points is smaller than distance between points between two different clusters. Hence, the choice of a distance metric for clustering is fundamental. The 'k-medoids' algorithm used in LCA studies, can work with both continuous and discrete data. Manhattan distance (eq. 2.10) is one of the widely used metric to measure the distance d between

two vectors [Wolfgang2011].

$$d_e(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n |x_i - y_i| \quad (2.10)$$

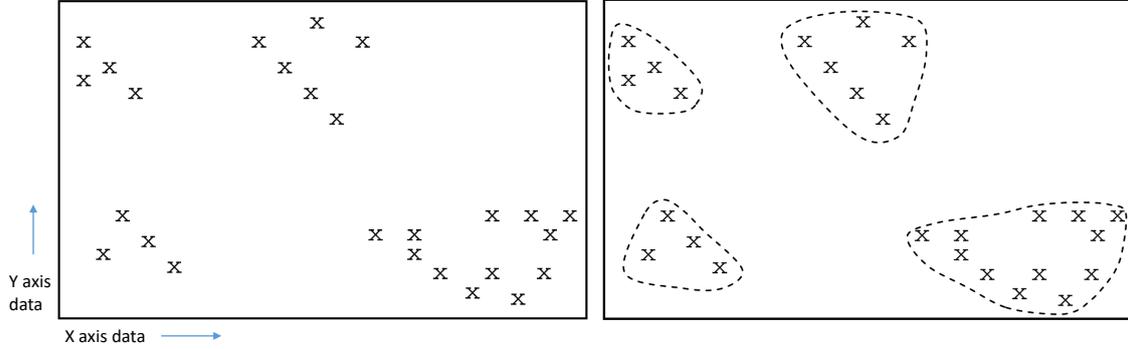


Figure 2.5: (Left) Objects for clustering, representation in a two dimensional data (Right) Clustering of objects [Bramer2007].

Classification is dividing objects into 'mutually exhaustive and exclusive categories', such that each object is assigned to at least one class and one object is assigned to precisely one class [Bramer2007]. The main difference here with clustering is that, in classification the data is labelled or already structured. Hence, the classification is categorised under 'supervised learning' while clustering is 'unsupervised'. Basic principles of a commonly used classifier: Naive bayes is described here with an example from [Skansi2018]. A fictitious record of whether the train was late according to the time of the day is presented in table 2.3. Here, some basic prior probabilities could be, such as train being late (9/13) and train being on time (4/13). Hence, to calculate $P_{on-time|morning}$, the probability of train being on time, given that it is morning will be:

$$P_{on-time|morning} = \frac{P_{morning|on-time} \cdot P_{on-time}}{P_{morning}} \quad (2.11)$$

Where, $P_{morning|on-time} = (2/4)$ is the probability that its morning given that the train is on time, while $P_{morning} = (5/13)$ and $P_{on-time} = (4/13)$ are the probabilities of it being morning and train being on time respectively. Thus supposedly for a new data point, $P_{on-time|morning} = 0.4$, thus train is predicted to be late or a value of 'yes'. This calculation can also be done directly from the table but it is used as a demonstration of the underlying principle. In practise, the value is calculated using much more complex data and probabilities. Thus, values of new data points are estimated by leveraging the prior collected data.

Application in LCA Sundaravaradan *et al.* performed clustering to group components and sub-components using the functions/text descriptions in the BOM and LCA databases [Sundaravaradan2011]. Then, the estimation of new data points was done using the other points belonging to the same cluster. They used this approach for 2 objectives: First, missing impact data of randomly selected processes from ecoinvent 2.0. Mixed results were obtained. High accuracies (>95%) for 90% of the points, while poor accuracies for the rest (up to 30%). Second, LCA calculation of new data points was done, by leveraging a BOM database of 560 components of printed circuit boards. The results are variable according to the object. The median accuracy was 87.2% and more than 75% accuracy for most of the objects, but high errors for the rest. Presumably when the supplied object is different to the dataset.

Hosain *et al.* also grouped objects in BOM and LCA database using clustering [Hossain2014]. Data was obtained from a computer manufacturer with 5,948 BOM descriptions. In the BOM

Table 2.3: Fictitious train data to demonstrate naive Bayes classifier [Skansi2018].

Time	Train late				
morning	no				
afternoon	yes				
evening	yes				
morning	yes				
morning	yes	Time	Late (yes)	On-time (no)	TOTAL
afternoon	yes	morning	3	2	5
evening	no	afternoon	4	1	5
evening	yes	evening	2	1	3
morning	no	TOTAL	9	4	13
afternoon	no				
afternoon	yes				
afternoon	yes				
morning	yes				

database, clusters were made according to the function satisfied using text descriptions. Meaning that the products present in the same cluster satisfy the same function and are 'substitutable'. While in the LCA database, clusters were made using text strings as well as their environmental impact. The differences between the two types of clustering (or disparate clustering) enabled automatic 'eco-design' recommendations. That is, sub-components of the objects providing the same functions but with a lower environmental impact. Furthermore, using the characteristics of the clusters, Bayes classifier is trained to map new activities in BOM database with the LCA database automatically. They achieved reasonably well matching of databases with majority of the matches having high confidence score (>0.95) and only 10% of the matches with a low score (<0.25). For the disparate clustering, a re-design of a desktop computer demonstrated upto 36% reduction in the carbon footprint.

Zhou *et al.* present a framework of performing LCA of buildings using data mining techniques [Zhou2015]. Their main recommendations are the following. First, using classifiers to map BOM activities to LCA databases. Second, for regional impact assessment of buildings, forming clusters of similar buildings in the region. LCA analysis only of small number of buildings as representatives of each cluster. Hence, regional impact could be found by scaling the clusters and thus indirectly, each building representative. Third, disparate clustering to propose design alternatives. However, no case study was presented to demonstrate their results.

Jeong *et al.* cluster the products utilising a 'Function-Behaviour-Structure-Environmental' characterisation of products [Jeong2014]. For example, a fan can be represented by: structural characteristics - steel, mass; behaviour - cross flow fan; function - move air; environmental - assembly by rolling and its environmental impact. Later, the similarities in the structure of components are used to find appropriate cluster to consequently calculate their impacts. They estimate pretty accurately (3-7% error) for 12 impact categories for case studies involving industrial fans. However, it should be noted that in this particular case there are only marginal differences between the training and the validation set.

Relevance Data mining techniques are meant to be used with large amounts of data generated from the rise of informatics. In terms of LCA, an important use could be mapping of external

databases with activities in LCA database. As it offered promising results in this case. This can reduce the effort required in manual mapping by an order of magnitude. Another useful application seems to be in 'eco-design' using disparate clustering on functional and environmental characteristics. As it can provide options satisfying the same function but with different environmental impact.

On the other hand, predicting impacts of new data points based on clustering similar products seems less attractive but could be a promising topic for future research. Two applications of this approach were: first, calculation of missing impact factors provided that other indicators for the same object are present. Moreover, the toxicity impact indicators are highly sensitive to the presence of certain compounds, which makes their prediction using this approach difficult. Second application is impact assessment of a new object. In this case, object to be assessed should be similar to ones already present. Even then it can only provide initial estimates. It can be argued that first estimates could be also made through a manual database or literature search.

Table 2.4: Recommendations on the use of the LCA modelling methods and relevance of their results

Approach	To be used when:	Relevance of LCA results
Parameterized LCA	<ul style="list-style-type: none"> - High accuracy is required - Detailed contribution analysis are required - Analysis often involve new technologies, processes and materials 	<ul style="list-style-type: none"> upto detailed LCA possible
Linear Regression	<ul style="list-style-type: none"> - LCA model is required in terms of an equation to simplify use or for utilisation with analytical methods - Prediction set is similar to the learning set 	screening LCA
ANNs	<ul style="list-style-type: none"> - Relationship between predictor variables and LCA impact is highly non-linear and predictor variables are correlated with each other - Comprehension of the model for contribution analysis is not required - Prediction set is similar to the learning set - Large amount of data is available 	<ul style="list-style-type: none"> Initial estimate or to fill data gaps for low importance data
Data mining techniques	<ul style="list-style-type: none"> - Not recommended for impact estimation, but 'classification' useful for mapping LCA databases with other databases - During that, 'Disparate clustering' could be useful in eco-design 	N.A.

2.2.5 Summary: Parameterized LCA models are selected but some issues remain unsolved

To satisfy the goals of this thesis, parameterized LCA models were selected as the computational approach for LCA models with the energy simulation software. Mainly since it was established in the context that often the modelling will deal with emerging energy technologies. In this case, the new predictions cannot be based on the existing ones since they could be remarkably different from each other. Furthermore, the objective of the environmental analysis is to make investment

and R&D decisions, which might require detailed analysis of sub-components of technologies. Hence, the level of detail into the sub-components offered by regression or ANNs are inadequate. Regarding data mining techniques, they could be useful when mapping is required between a large technical database with an LCA database, which is not the case currently. Recommendations on when to use these strategies for integrating with external models are presented in table 2.4.

However, there are two problems associated with development of parameterized LCA models in this context:

First, is the complex nature of modelling. It requires high level of LCA expertise, good knowledge of the technology, insights on its future evolution and familiarity with energy simulations. Moreover, the modelling is arguably more complex as compared to the conventional LCA studies. Since all the outputs from energy simulations influencing the LCA indicators have to be taken into account. On the other hand from the LCA perspective, modelling of inventories of components, their future evolution, variation due to geography and end of life treatment is not taken into account in the techno-economic simulations. Hence, in addition to the technical parameters, new parameters pertinent to LCA might have to be introduced.

Second, owing to these variabilities, modelling becomes time intensive. During the design of hybrid energy systems for the future, there are often many uncertainties related to design parameters and background/foreground processes. As a result, prioritizing the data collection becomes indispensable. Due to the complex nature of the simulations, multiple uncertainties and their possible co-relations, local sensitivity analysis as practiced in conventional LCA studies might be inadequate and tiresome.

It appears that the techniques and strategies to solve the above issues do exist in the LCA literature. However, a consistent methodology that combines them to enable the establishment of parameterized LCA models for hybrid energy systems does not exist yet.

2.3 Inclusion of LCA indicators in optimisation

In this section, the problem of inclusion of LCA indicators in optimisation is confronted with the techniques in the literature. The importance of this inclusion was justified earlier. To reiterate, it will enable users to search for solutions with higher environmental efficiency. Thus the problems that usually involve deciding the system configuration or the operational strategy, could be viewed from an additional perspective of environmental impacts.

2.3.1 Problems with the currently used genetic algorithm SPEA2

As mentioned earlier, SPEA2 [Zitzler2001] is currently utilised in Odyssey for searching solutions with respect to maximum 3 objectives simultaneously. It is one of the landmark 2nd generation evolutionary algorithms [Bechikh2016]. However, numerous investigations have shown that its performance considerably deteriorates when number of objectives increase more than 3 [Li2015, Li2018, Deb2013]. The same is also true for its contemporaries (e.g. performance of NSGA-II demonstrated by [Saxena2012]). In this thesis, at least one technical and economic indicator is expected to be present. Furthermore, a holistic perspective of environmental impact is envisaged, thus it is likely that number of objective increase more than 3. Continuation of the same algorithm can thus become problematic.

The main reason for the inefficacy is that when the number of objectives increase, more solutions become indistinguishable from each other. That is, the number of non-dominated solutions for selection increase exponentially with objectives [Li2015]. Recall figure 1.3, after fitness evaluation, the problem appears when solutions have to be selected so that size N becomes equal to

the archive size S . Now, the secondary selection operator based on diversity is deployed. The algorithm is therefore unable to adequately push the solutions towards convergence. As a result of this phenomenon, the final solutions might represent good diversity without converging towards the pareto front [Li2015, Li2018, Saxena2012].

2.3.2 Use of genetic algorithm NSGA-III in this thesis

There are numerous approaches for solving multi-criteria optimisation and even more algorithms. For example, [Marler2004] identified 21 different approaches including evolutionary algorithms for many-objective optimisation. While [Li2018] identified 13 different, equally well performing evolutionary algorithms only. They are currently preferred in the laboratory for the searching solutions. They are relatively robust and can handle variety of different problem types: non-linear, continuous/discrete and even black-box models [Marler2004]. Another advantage is that its computations can be run in parallel. Thus multiple computers can be deployed to solve a large problem. Given the experience of the lab members (both users and tool developers) with evolutionary algorithms and the past research [Nadal2019, Guinot2015b, Guinot2015a], this approach is kept the same. Additionally, it will be easier for the users to process additional information concerning the impact with already utilised approaches for dimensioning. It also facilitated the implementation of the genetic algorithms during this thesis.

Another important point is that LCA calculations depend on the energy simulation models in the first place. Thus it seems logical to integrate LCA indicators in the optimisation approaches similar to the ones already selected for energy system optimisation.

Amongst the evolutionary algorithms, NSGA-III was selected as it is considered as one of the state-of-the art evolutionary algorithms for many-objective optimisation [Li2018]. It was first introduced in 2014, in response to the convergence problems mentioned in the previous subsection [Deb2013]. Its performance in terms of convergence and diversity for upto 15 objectives was also demonstrated [Jain2013]. Another key advantage of this algorithm is the preferential search of the pareto front. Due to its reference points based approach, it inherently searches for preferred points. It becomes convenient since with high number of objectives size of the pareto front becomes large and not all parts could be of interest. As a result, the post-pareto processing to select particular solutions for decision-making might not be necessary.

Description of NSGA-III

A simplified illustration of the algorithm is provided in figure 2.6, it is described in detail below. The description is provided using the original paper [Deb2013] and its implementation in this thesis using the DEAP framework in Python [Fortin2012]. The first step is the specification of reference points which essentially dictates the preferred pareto points to be searched. Then a pseudo random population is initialised. Meaning random individuals (system configurations) within the search range. For example, for solar PV size range of 0-8MW, random sizes between this range are generated for the individuals in the first generation. Often the population size is slightly higher than the number of reference points. Offsprings from the parent population are produced by applying crossover and mutation operators. Fitness or the performance indicators of each individual in the two populations are then evaluated. Based on them and the reference points, population for the next generation is selected. The approximation of the pareto front is updated at each generation using the population by retaining only the non-dominated members. The population size remains constant but the number of pareto front members likely increase with generations. These steps are described in detail below:

Reference point specification Reference points in NSGA-III are specified in terms of the objectives either as a uniform spread or as a set of preferential points. These points lie between 0

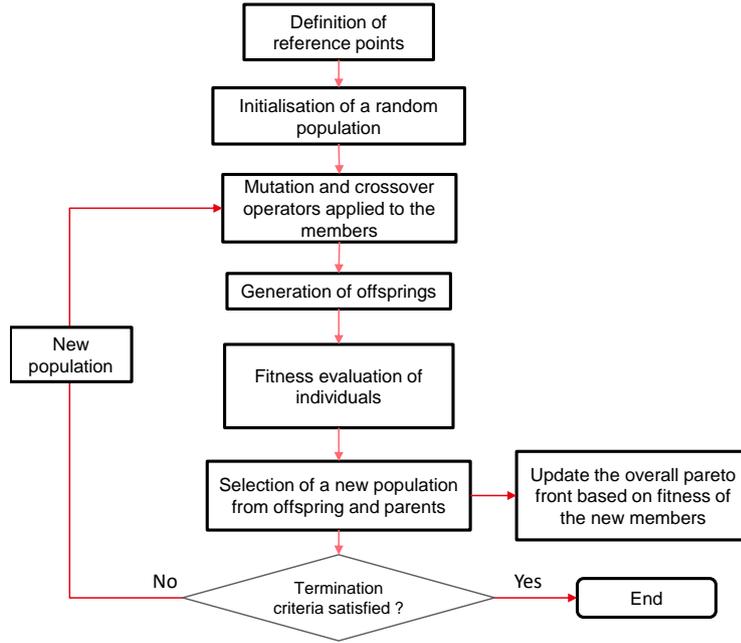


Figure 2.6: NSGA-III operating scheme in this thesis.

to 1, representing the min and max value of each objective respectively. These points enable the convergence and diversity of solutions. Furthermore, since they are associated with the proposed solutions they allow preferential search of the pareto front. As the number of objectives increase, the size of the pareto front increases exponentially. For decision-making, we might be interested only in a part of the Pareto front or in having a good distribution of points across the entire Pareto front. In this context, the use of reference points thus becomes interesting. Another advantage in this aspect is that the use of processing of pareto front using techniques such as multi-criteria decision making might not be necessary. Since the preferential solutions in pareto are inherently searched.

To obtain a uniform spread, approach from [Das1998] is proposed by the authors of NSGA-III. It involves placing points on a normalized hyperplane with an intercept = 1 on each axis. The number of points (H) is calculated as:

$$H = \frac{(N_{OBJ} + \hat{P} - 1)!}{\hat{P}! \times (N_{OBJ} - 1)!} \quad (2.12)$$

Where, \hat{P} is the number of axis partitions and N_{OBJ} is the number of objectives. These reference points increase exponentially with the number of objectives. For example, for 3 objectives with 10 axis partitions, 66 reference points are obtained shown on the left in figure 2.7. Similarly, other preferential set of points can also be decided. In this case it should be noted that extreme points at $\hat{P}=1$ and scale=1 also have to be provided.

Whereas when the number of objectives are increased to 10, number of reference points become 92,378. To keep this in check, the authors propose decreasing the no. of axis partitions but using using multiple layers at different scales. For example on the right in the same figure, layering approach in 3 dimensions is shown when the objectives increase.

Initialization of a pseudo-random population. In the first step, an initial set of individuals are specified using random values within the individual search space. An individual is the set of values to be optimised. For each value, a search limit is to be provided. This can be based on engineering

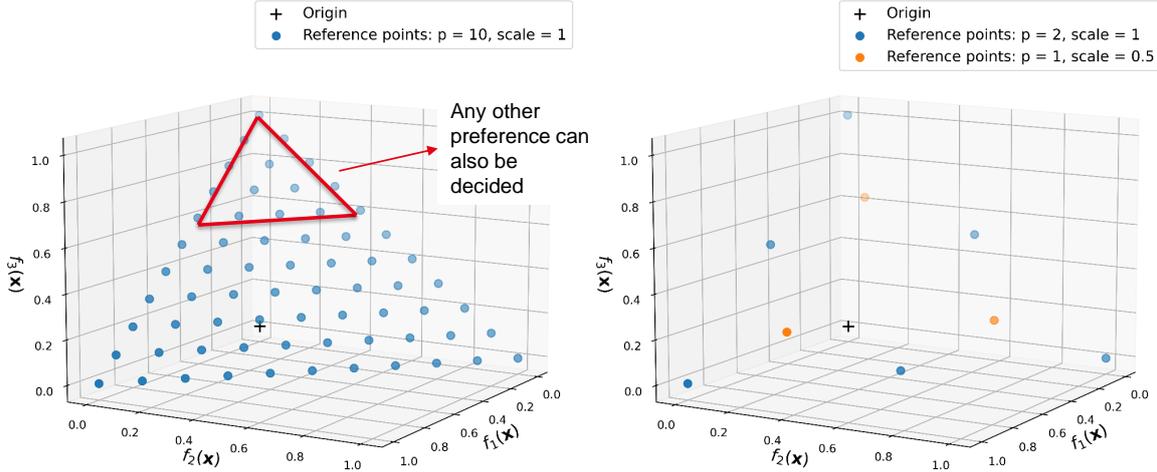


Figure 2.7: Reference point distribution illustrated for 3 dimensions. (Left) for 10 axis partitions. (Right) using a layering approach when the number of objectives increase.

estimations. A set of individuals then constitute the initial population. The population should be large enough to have a good distribution of individuals in the 0^{th} , otherwise there can be a risk of the optimisation getting stuck at a local optimum.

Selection This step is probably the most complex part of the algorithm. Once the offsprings are created, selection of population for the next generation is made on the basis of fitness from the offsprings and the parents. The latter can also be referenced as the current generation individuals.

Ranking the population into levels of domination. When selecting a set of population for the next generation, the members are distinguished in non-domination levels (F). That is, each non-domination level consists of individuals that are indistinguishable from each other (or non-dominated). Individuals in the lower levels are dominated by the ones in the level above. When the selection step proceeds, members are filled from the first non-dominated level F_1 . In simple terms, the non-dominated solutions are filled first then the algorithm goes to the second best level until the number of selected members is equal to or more than the population size. If the size of population is equal to the number of individuals in the selected levels F_{lev} , no further steps are required and the next generation can be started. However, often this is not the case especially when the number of objectives increase and large number of individuals become non-dominated. Selection is to be made amongst individuals at a particular level F_{lev} such that the population size remains constant throughout the evolutionary process. For this, individuals are differentiated according to their diversity using the reference points in the following steps. This is done by ensuring that the population represents the desired pareto solutions. All other solutions in the levels beyond $F_{(lev+1,2,..)}$ are discarded.

Standardization of fitness. In this step, the fitness of each individual is standardized with respect to the rest of the population. First the ideal (z_{ob}) and extreme points (a_{ob}) of the selected population are defined by assembling the minimum and the maximum fitness value respectively of each respective objective. Supposing $f_{ob}(x)$ is the fitness value for an objective ob , the standardised value of the fitness can be given by:

$$f'_{ob}(x) = \frac{f_{ob}(x) - z_{ob}}{a_{ob}} \quad \text{For } ob = 1, 2, 3 \dots N_{OBJ} \quad (2.13)$$

This standardisation procedure facilitates dealing with problems when the objectives are differently scaled and lie in different orders of magnitude. Which is often the case across different LCA impact indicators.

Association of reference points with the population. The standardised points in the population can now be associated with the reference points defined earlier. It is interesting to note that after the standardisation, these points will lie on the same plane as the reference points. For association, a line is drawn from the origin to each of reference points. Then from this line, the perpendicular distance to the standardised points is calculated. The individual with the shortest distance to the reference point, is assigned to this point. A reference point can thus have multiple or zero individuals associated to it. This procedure of association of individuals before standardisation, is illustrated in figure 2.8.

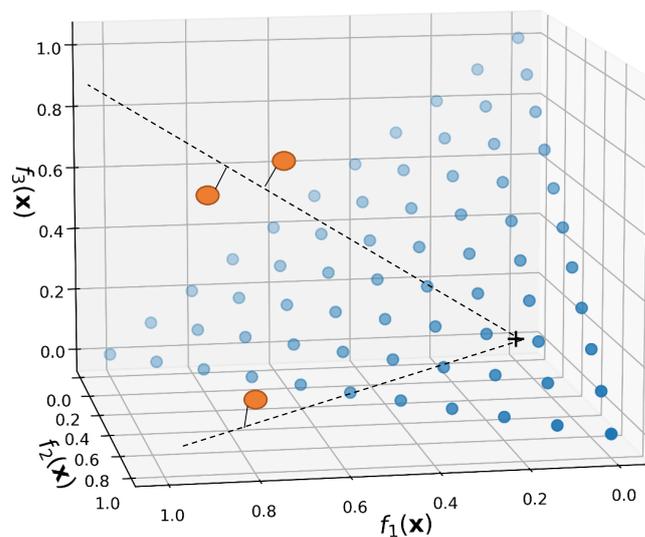


Figure 2.8: Illustration of the association operation of individuals (orange points) with reference points (in blue) [Deb2013].

Niche preservation. To ease the explanation of this step, it is re-iterated that individuals in $F_{(lev-1)}$ levels are already selected. On the other hand, the level F_{lev} has too many members for selection. A niche count (ρ_j) is defined as the number of already selected individuals (in levels $F_{(lev-1)}$) associated with a particular reference point (j). The first step is to identify reference points with minimum ρ_j value. If $\rho_j=0$, and if this reference point has no associated members in either in levels $F_{(lev-1)}$ or F_{lev} , this point is discarded from further evaluations until the next generation. Or else if there are members of F_l associated with this point, one with the shortest distance as calculated previously is added and the niche count is increased by one. Whereas when $\rho_j \geq 1$, a member randomly selected amongst the ones associated with this point from front F_{lev} is added to the next generation. Leading to an increase in ρ_j . All the vacant population slots are thus filled in this manner.

Optimisation loop Once the new generation is selected, offsprings are again created using mutation and mating operations and the loop proceeds until the stopping criterion is satisfied. Examples of stopping criterion can be (lack of) change in objectives for a number of successive generations, analysing the optimisation indicators (e.g. hypervolume or IGD) or pre-defined number of generations.

2.3.3 Optimisation overview in life cycle assessment

In order to search for solutions while including the LCA indicators, multi-objective optimisation has been widely used. Often this process included LCA indicators with other economic or technical indicators. Owing to the added complexities with higher objectives in optimisation, often only a limited number of them are included. At the output, pareto solutions are processed further or directly presented for decision-making.

Amongst them, two approaches were commonly found in the LCA literature were ϵ -constraint method and genetic algorithms. It should be noted that a variety of other approaches are possible and available in the optimisation literature but these two were often employed.

ϵ -constraint method is one of the commonly used methods for multi-objective optimisation. For this method initially all objectives are separately optimised to obtain a minimum value for each of them. Then, a primary objective is identified which is then optimised but in this instance, placing constraints on the secondary objectives. Constraints are modified (relaxed or tightened) for each optimisation run such that they cover the entire optimisation space in the desired steps, while placing constraints on the secondary objectives [Vandepaer2020].

The earliest examples including the ϵ -constraint method are found in the late 1990's where [Azapagic1999] used this approach to simultaneously optimise production, costs and global warming potential in order to manage a chemical process chain. [Zhang2015] also used ϵ -constraint method to assess various system configurations in micro-grids. They included global warming potential and acidification potential as two environmental indicators. [Yue2016] proposed a hybrid LCA, multi-optimisation framework to include green house gas emissions in the decision-making of biomass supply chains in the UK. [Antipova2014] include 6 environmental indicators with other conventionally used indicators such as cost to investigate alternatives for retro-fitting of buildings. More recently, [Vandepaer2020] include 3 indicators along with costs to investigate future energy scenarios in Switzerland.

Genetic algorithms as introduced earlier have also been used to include LCA indicators as optimisation objectives. Some examples are presented here. [Gerber2011] used a genetic algorithm in the process design of a biomass plant utilising wood chips to generate synthetic natural gas and heat/electricity. They used thermodynamic design as the technical constraint, costs and Ecoindicator99 as an aggregated environmental indicator to search acceptable solutions. [Ahmadi2015] used NSGA-II, a pareto based genetic algorithm for eco design of conventional water production process. Three objectives were included in the main optimisation loop for this purpose: water quality indicator, costs and an aggregated environmental score. Design of hybrid energy systems, similar to the ones in thesis were done by [Nagapurkar2019] using genetic algorithms. They included CO₂ emissions with the techno-economic design of microgrids in United States. Sizing and operation of pharmaceutical batch plants were optimised by [Dietz2006] while using genetic algorithms. They included an aggregated environmental impact and costs as the optimisation criteria.

Problems with limited LCA indicators

It was seen that although LCA indicators have been included in optimisation, only a limited number of them are usually included. Either they are included in an aggregated form or CO₂ emissions is taken as a proxy for all environmental impacts. When the former is done using weighting or endpoints, there are two problems. First, there is a risk of ruling out solutions that are optimal in the original objective space. Since the objectives disproportionately contribute to the final impact, the dominance structure of the original problem is distorted. This was clearly demonstrated by [Antipova2013], they found that using only the total Ecoindicator-99 favoured only one group of indicators. It resulted in optimisation algorithm omitting one set of otherwise optimal solutions.

The second problem is the uncertainties or the subjective choices involved in the weighting and endpoint indicators. LCA community as a whole does not seem to favour the use of these two

approaches possibly due to the lack of robust conclusions that could be drawn from them. More discussion on it in section 4.1.2.

On the other hand when selecting only a limited set of indicators without taking into account the pareto structure of the problem might lead to impact transfer. The two reasons outlined by [Guillén-Gosálbez2011]. First, we might have an impression that solution A is superior to solution B since it has lower impacts in the selected indicators. However, solution B could be better than solution A in the omitted objectives. Secondly, two solutions can have same performance in the reduced objective space while one of them could be superior when viewed with all objectives.

Impact transfer is even more important for the emerging technologies. It could happen that designers or engineers are not aware of an impact that could become problematic in the future. On the other hand, the problematic aspects could be addressed if they are identified during the development.

2.3.4 Objective reduction techniques in life cycle assessment

It should also be noted that including large number of objectives in optimisation is not straightforward as it leads to various issues such as high computational cost, convergence problems of algorithms, visualization and expensive calculation of algorithm performance metrics [Deb2013, Saxena2012]. One of the methods to alleviate some of these problems, is by temporarily eliminating the redundant objectives [Li2018]. They rely on the fact that when a many-objective problem has the same pareto front with fewer objectives, optimisation can be made on the lower objectives instead of the original ones [Li2015]. These techniques are quite relevant in life cycle assessment where the indicators are often found to be correlated.

There are various objective reduction techniques in the literature that could be classified under two types: *offline* and *online*. In the *online* methods, the objective reduction is performed gradually as the algorithm progresses. While in the *offline* methods, objective reduction is done after obtaining a set of solutions. In LCA investigations, three methods of this type were found.

A general framework of the objective reduction methods utilised in LCA can be seen in figure 2.9. First step is the partial initialisation to obtain a sample set. Partial initialisation is optimisation with respect to arbitrary reduced set of objectives or optimisation in the complete space of tech-eco-env objectives but with reduced amount of time. The objective reduction techniques then analyse the initial set of solutions to select a reduced set of indicators. The optimisation algorithm then works with respect to the selected indicators.

Two types of objective reduction techniques were found in LCA: Correlation based and dominance structure based. In the latter, correlations between the objectives are leveraged. While in the former, redundant objectives are identified using the preservation of the dominance structure between the solutions [Li2015]. Yuan *et al.* performed a comprehensive analysis of objective reduction methods and found that none of the two is clearly superior [Yuan2017]. Both have their respective advantages and limitations. Three objective reduction techniques were found in LCA. Out of them, the first two are based on correlations while the last one is based on dominance structure preservation. They are described as follows:

Principal component analysis (PCA) is a statistical tool that is used for analysis of complex data. More specifically, identification of redundant variables that do not add new information but complicate the data. It allows the transformation of co-related variables into a set of ordered uncorrelated variables, known as principal components (PC). PCs can be ranked according to their ability to explain the variance in the data set [Deb2005]. In many practical problems only a few PCs explain most of the variance in the entire set.

Objective reduction based on PCA was first proposed by [Deb2005]. They then provided further customization to the reduction framework in [Saxena2012] for different types of problems. In the context of LCA, each PC is a combination of original impact indicators. Once the PCs

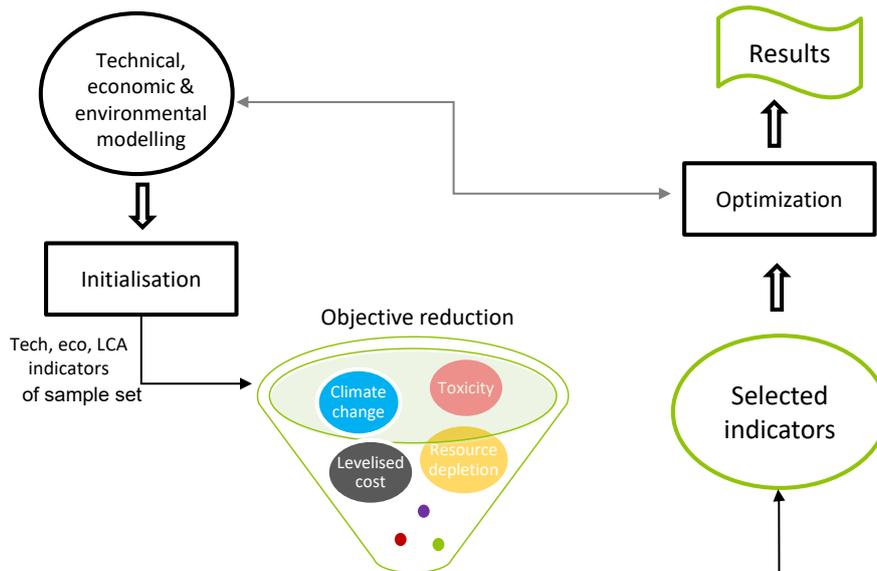


Figure 2.9: General objective reduction scheme used for optimisation including LCA indicators

explaining the maximum variance are identified, indicators making the maximum contribution in them can be selected.

The following are the steps to calculate PCs and to reduce the number of objectives are described as per [Deb2005, Saxena2012]. First step is the standardisation of the obtained dataset by subtracting mean of each indicator from each datapoint. From this dataset, a correlation matrix is then calculated for the LCA indicators. Eigenvectors and eigenvalues of the matrix are then calculated. The first PC is then the eigenvector corresponding to the highest eigenvalue. Second PC is the eigenvector with the second highest eigenvalue and so on. Next, only the PCs are retained which explain the maximum variance. The generally recommended value is 99.7% but for problems with high number of redundant indicators, [Saxena2012] demonstrated that even a values as low as $\approx 68\%$ could be used that explain most of the variance in the original objectives.

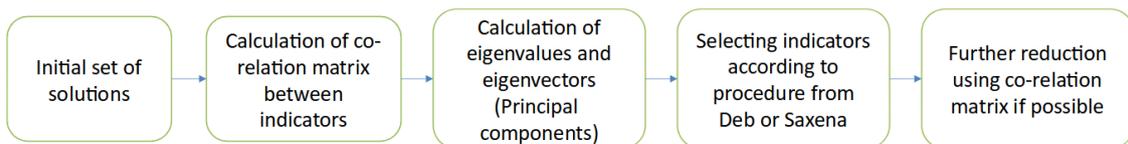


Figure 2.10: An illustration of the methodology to select indicators using PCA heuristics by [Deb2005, Saxena2012].

Impact indicators in the retained PCs are selected according to two rules:

1. Objectives with most negative and most positive values are chosen
2. If all objectives have the same sign, two objectives with the highest values are chosen

Then in the final step, the selected indicators are further reduced according to correlation coefficients between them. For this step, [Saxena2012] introduced quantifiable parameters to facilitate this step such as correlation threshold and a selection score.

Life cycle assessment can also be considered as a highly redundant problem if large number of indicators are considered. Since it was found that only 6 indicators explain 92% of the variance out of a total of 135 impact indicators for 976 products using PCA [Steinmann2016].

Application in LCA Gutierrez *et al.* were the first ones to use PCA in LCA [Gutiérrez2010]. They used it to gain insights on the environmental impact indicators of household products. More specifically, they used PCA and other similar techniques to find correlations between indicators. Thus, a small subset of 3 indicators could explain relative performance of products instead of the total 12 utilised.

Sabio *et al.* were the first ones to utilise PCA specifically for LCA objective reduction in optimisation [Sabio2012]. They applied the strategy to design hydrogen supply chains in Spain. The problem was formulated as a mixed integer linear programming (MILP) problem followed by ϵ -constraint optimisation. They were able to discard 4 out of 8 initial indicators. Similar problem formulation was done by [Poza2012], who applied this method to 2 case studies. Brunet *et al.* also followed a similar approach of MILP formulation, PCA and ϵ -constraint optimisation to design bioprocesses (chemicals) [Brunet2012]. More recently, [Perez Gallardo2013] applied PCA to reduce objectives and then to dimension a PV system using genetic optimisation.

Representative objective reduction method (ROM) This method proposed by [Cucek2013], relies on determination of correlations between the indicators and then selecting a representative set amongst them. After obtaining pareto-optimal points with respect to all objectives, these points are standardized by dividing each one by the maximum of each indicator. A handful of (5-6) equidistant points on the mix-max range of each objective are then selected. Three values are then calculated between all possible pairs of indicators, in order to group them:

- Normalised ratios between pairs of footprints (should be close to 1)
- Overlapping pairs of footprints within the process variables (should be close to 1)
- Average absolute normalised deviation between pairs of footprints (should be close to 0)

One indicator each from the uncorrelated sets formed is then selected for optimisation. Cucek *et al.* optimised the integrated economic-environmental model using the ϵ -constraint method [Cucek2013]. They then utilised ROM to reduce 5 environmental indicators to 3. A less rigorous and more practical application of ROM was done in [Čuček2014] also using the biomass supply chains as a case study.

Dominance structure based reduction Dominance structure based technique for objective reduction was first introduced by [Brockhoff2006]. It was applied to LCA for the first time by [Guillén-Gosálbez2011]. It relies on omitting indicators such that the rankings of the solutions are unchanged. For example, if we consider solutions A and B with three objectives f_1 , f_2 and f_3 in figure 2.11. If all objectives are to be minimised, A and B are indistinguishable, since neither of the two is better in all of the three objectives. Now, if say objective f_3 is omitted, solution B becomes superior to A with an error of $\delta=0.4$. On the other hand when objective f_2 is omitted, A and B remain indistinguishable i.e. the initial Pareto structure of the problem is maintained even after omitting one objective. This technique thus relies on MILP based formulation to discard indicators such that the error ' δ ' is minimised.

Application in LCA After the initial introduction in [Guillén-Gosálbez2011], slight modifications or new metrics on the same principles were proposed by [Vázquez2018]. Redundant environmental indicators were identified during dimensioning of a solar reverse osmosis plant [Antipova2013]. They identified two sets of indicators which followed the same trends. They use the same case study to further propose use of filters for *post-optimisation* analysis [Antipova2015].

Hennen *et al.* propose a objective reduction framework to dimension energy systems based on the same approach [Hennen2017]. They applied their framework to dimensioning problem in

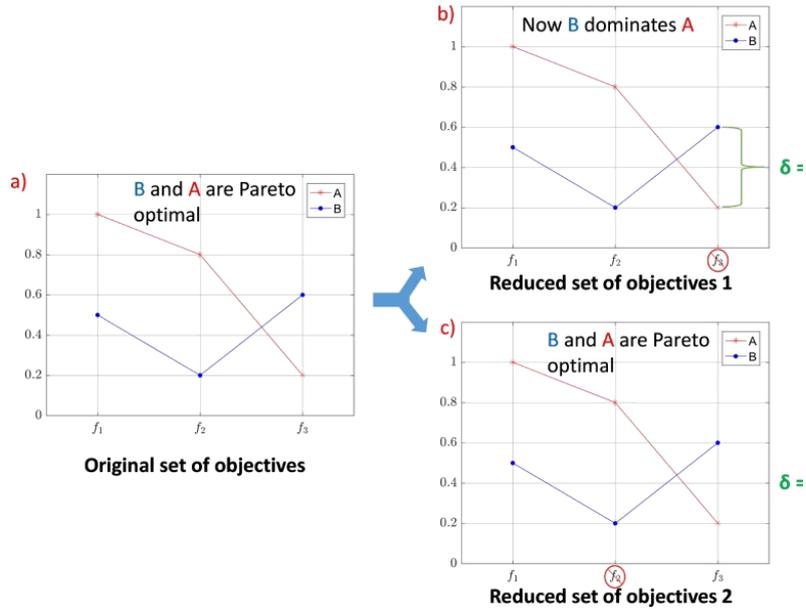


Figure 2.11: An example to demonstrate dominance-structure based reduction technique. (a) Three original objectives (b) Objective f_3 discarded (b) Objective f_2 discarded. Illustration from [Vázquez2018].

heating and cooling networks operating on natural gas and electricity grid. Kostin *et al.* also use objective reduction to design sugar cane supply chains in Argentina [Kostin2012].

LCA was integrated with specialised chemical process design software for an integrated design including economic, technical and environmental criteria by [Gonzalez-Garay2018]. Along with the reduction of environmental objectives, they used ANNs for approximation of their integrated model, data envelopment analysis as a post filter of solutions and hypervolume indicator to verify the quality of pareto front.

2.3.5 Summary: A possible combination of NSGA-III with objective reduction?

NSGA-III appears to be a sound algorithm for the many-objective optimisation problems involving life cycle assessment that appear in this thesis. Its combination with objective reduction techniques could have two interesting benefits:

Computational time and quality of results The peculiarity in the case of hybrid energy simulation software is that they are often computationally expensive. For instance, the a single simulation of case study in figure 1.10 takes around 9 seconds to run on a single processor. If for instance say uniform solution distribution for a 15 objective problem is desired, the number of reference points for $\hat{P}=2$ at 2 scales become 240 according to equation 2.12. Population size which is the next higher multiple of 4 then becomes 244. If the simulation time is calculated for parallel computing on 8 processors, it is estimated to be around 22.9 hours for 300 generations.

The life cycle impact indicators are often found to be correlated. The vast number of studies found in the previous subsection are an evidence of this. If the redundant indicators in LCA could be identified and it appears that the performance of NSGA-III could be further improved.

Let's say in a hypothetical scenario, 15 indicators are reduced to 6 using one of the reduction techniques. Which is not a far-fetched assumption given the past studies [Steinmann2016, Perez-Gallardo2018]. For 6 indicators, using $\hat{P}=3$ at 2 scales, the number of reference points become 112

and the population size =116. The simulation time is then reduced to 10.9 hours, less than half the previous time. This is mainly due to the reduction in the number of simulations required. There is another advantage to this, the increase in number of axis partitions (\hat{P}) at both scales mean that more distributed solutions can be obtained. To summarize, it appears that higher quality solutions could be obtained in less than half the time.

Visualization A second advantage can appear if correlation based objective reduction techniques are used. If groups can be made of indicators that co-vary, the visualization in the final set of solutions can be greatly improved. For instance, only the reduced set of indicators could be provided to facilitate decision-making along with the correlation information of omitted indicators. As mentioned earlier, LCA could be considered as a redundant problem. Consequently, a small set of indicators could be sufficient to explain the changes in rest of the indicators. This could significantly improve visualization even if large number of LCA indicators are considered in the initial problem.

Thus a combination of NSGA-III with objective reduction techniques seems promising. However, it is not yet done in the literature.

2.4 Synthesis of research proposals

In order to facilitate the consideration of environmental impacts in the assessments of hybrid energy systems, the two major goals of the thesis were established at the end of chapter 1. These goals being: integrated tech-eco-LCA modelling and inclusion of LCA indicators in the search of feasible solutions for hybrid energy systems. They were then confronted with the state-of-the-art in sections 2.1, 2.2 and 2.3, to formulate three research proposals of this thesis.

Proposal 1. At the end of section 2.1, it was found that the problems regarding appropriate consideration of environmental impacts also exist in other available software that analyse hybrid energy systems. Then after looking at the existing approaches that combine LCA computations with external tools, it was found that an approach defined as parameterised LCA modelling could be leveraged for integrated tech-eco-LCA modelling. However, a systematic methodology is missing that takes into account various problematic aspects in this modelling context. These aspects refer to variabilities in geography, scaling, data gaps and future evolution. Another element that complicates these elements further is the presence of emerging technologies. Additionally since this methodology could be used by personnel beginners in LCA to develop LCA models for hybrid energy systems. Thus a method to reliably and effectively develop these models are necessary but missing so far in the literature. This proposal is thus explored in chapter 3.

Proposal 2. Then from section 2.3 it was concluded that a new generation evolutionary algorithm, NSGA-III could be used to include LCA indicators in the optimisation process. It was also found that utilising it in combination with objective reduction techniques might offer improvements on its performance in terms of computational time and quality of solutions. To the best of our knowledge, NSGA-III has not been combined with objective reduction thus far in the literature. Also, it appears that this algorithm has not yet been used with LCA indicators. This proposed approach is thus tested in chapter 5.

Proposal 3. During the implementation of the algorithm in the above proposal, it was realised that correlations between LCA indicators could be identified without generating an initial sample

set. As a result, redundant objectives could be reduced without an initialization of the problem. As far as it is understood by us, this is not possible in other reduction techniques in the literature. It has limited application than other techniques but it can result in saving computational time and uncertainties related to the generation of the sample set. This proposal is described and demonstrated in [chapter 4](#).

3

Establishment of parameterized life cycle assessment models for hybrid energy simulations

In this chapter, the first proposal of this thesis is elaborated, which is, framework for the establishment of parameterized life cycle assessment models of hybrid energy systems. The first section is commenced by looking at other similar investigations in the literature and the contribution of this work amongst them. Then some background on the approach followed for the development of this chapter and its relevance is described. Then in the second section, the framework is described, and it is implemented in the next section using a PEM water electrolysis case study in hybrid energy simulations. The results of the case study application are seen in the fourth section. Finally, the chapter is concluded in the last section.

Contents

3.1	Introduction	53
3.1.1	Contribution to the existing literature	53
3.1.2	Framework development approach	53
3.1.3	Relevance of the framework	54
3.2	Framework to establish parameterized LCA models for hybrid energy simulations	54
3.2.1	Handling inputs from energy simulations	55
3.2.2	LCA modelling approach	55
3.2.3	Dealing with variabilities and uncertainties	57
3.2.4	Model assembly and parameter definition	61
3.2.5	GSA using Derivative-based Global Sensitivity Measure	63
3.3	Case study: PEM water electrolysis	66
3.3.1	Handling outputs from the energy simulation	67
3.3.2	LCA modelling approach	68
3.3.3	Addressing variabilities and uncertainties	69
3.3.4	Model Assembly	70
3.3.5	Parameter definition for GSA	71
3.4	Case study: Results	72

3.4.1	Sensitivity & Importance indices	72
3.4.2	Improving the influential aspects for the scope	73
3.4.3	Results for scope 2030 and comparison with ideal case & literature values	74
3.4.4	When influential parameters cannot be fixed	75
3.5	Conclusions	76

3.1 Introduction

3.1.1 Contribution to the existing literature

The relevance of parameterized LCA modelling approach for the purpose of integrating LCA with energy simulations was justified in section 2.2.5. It will facilitate the estimation of LCA indicators with the corresponding alterations in the energy simulations.

Nevertheless, two problems concerning the complex and resource intensive nature of modelling, appear during the development of these models for hybrid energy simulations. The modelling is complex since a good expertise of energy simulations and technologies is required. Secondly, due the inherent variabilities of energy technologies and additional uncertainties because of the emerging technology context, the modelling requires different types of data. Thus it becomes necessary that the data to be searched is assigned relative importance. In its absence, data which might not have significant influence on the results might be searched preferentially. Furthermore, since the users in the laboratory are only beginners in LCA, many issues such as filling data gaps, scaling inventories, accounting for future context, etc. might not be clear for them.

If we look at the literature, multiple parameterized LCA models with varying levels of detail were identified in section 2.2.4 such as to aid decision-making of cranes [Ostad-Ahmad-Ghorabi2011], packaging [Verghese2010], milk production [Pirlo2019], electric motor vehicles [Nordelöf2018] and PV panels [Gazbour2019]. Gerber and colleagues present a systematic methodology for LCA integration with process design framework [Gerber2012, Gerber2011]. They indeed present important work in terms of LCA integration and scaling of equipment. However, they do not address the issues related to future evolution of systems and guiding data collection due to multiple uncertain factors.

Establishment of parameterized models followed by identification of influential parameters using global sensitivity analysis is applied in the following contexts for different reasons: wind turbines [Padey2013], bio-refineries [Pérez-López2018], buildings [Mastrucci2017] and geothermal energy [Lacirignola2017]. In this thesis, these studies are built upon to propose a framework to establish parameterized LCA models for hybrid energy systems. These models are particularly suited for *ex-ante* LCA of emerging energy technologies and integration with hybrid energy simulations.

The second goal of the proposal is to save time in modelling through prioritising data collection of parameters and LCA processes. Influential parameters are recognized using derivative based global sensitivity measure (DGSM). While the influential LCA foreground processes are identified using a novel coefficient, importance index, based on DGSM. This global sensitivity analysis method, DGSM, has not been used in LCA studies yet.

The resulting LCA model at the end of the framework is such that, it provides users with LCA indicators as the values or design in the energy simulation change. It thus also allows its use in tools such as optimisation that are often used to analyse energy systems. Consequently enabling the consideration of LCA indicators in the investigations of hybrid energy systems. Thereafter, the methodology is demonstrated using a case study of a PEM electrolyser. We show that only improving the identified important aspects of the model is sufficient to obtain reasonably accurate results.

3.1.2 Framework development approach

The first step towards developing the framework were the discussions with the users in the laboratory and utilisation of energy simulation tools. The goal was to follow the analysis process of energy systems, outline ideal requirements from LCA integration model and note the issues faced by engineers during modeling. The first chapter was partially written on the basis of these issues.

This was followed by comparison of hybrid energy simulation software, the exercise presented in section 2.1. An overview of calculation models for environmental models with energy simula-

tions was thus obtained. Then, the LCA computation models for integration with different tools in the literature were reviewed in section 2.2. Finally, to find useful strategies for the emerging technologies, references collected by the following studies on ‘ex-ante LCA’ were investigated and relevant ones were added [Arvidsson2018, Thomassen2019, Moni2020, Bergerson2020, Buyle2019].

Once a preliminary methodology was prepared, it was given to an engineer to develop the model of PEM electrolyser. Based on the feedback from the engineer, the methodology was improved which is finally presented here.

3.1.3 Relevance of the framework

The framework is primarily aimed at personnel developing LCA models for hybrid energy simulations: energy system engineers, software developers or LCA practitioners. It is based on the best practices found in this context. It is assumed that the personnel using this framework are beginners in LCA with at least a basic training. This framework does not aim at replacing the standard LCA methodology [European Commission2010] but to facilitate its implementation in the context of hybrid energy simulations. More specifically, it aims to support readers in the following steps of effectively integrating LCA with hybrid energy simulations: inventory analysis, impact assessment and interpretation. Thus, the framework comes into play once the goal and scope for the LCA investigation is finalised.

3.2 Framework to establish parameterized LCA models for hybrid energy simulations

A schematic briefly representing the methodology is presented in figure 3.1, where the steps proposed in this chapter are outlined in black. The goal & scope is defined as per the the steps of an LCA [European Commission2010], for a hybrid energy simulation model. The next step is the starting point of the framework, where guidance is provided to facilitate recognition and refining outputs from energy simulations for LCA calculations. Then, for the identified processes, level of detail for LCA models is decided. This step also includes steps of building inventory for emerging technologies if the data is not readily available. The next step is instructions on dealing with the uncertainties and inherent variabilities of energy components. Then recommended assembly of LCA models with respect to the energy simulations is presented along with definition of model parameters. Global sensitivity analysis (GSA) is conducted using DGSM to identify influential parameters and LCA processes. They are then improved in an iterative manner to obtain an LCA model integrated with energy simulations.

In essence, the methodology involves building the model with easily available data but ensuring that as much as possible all the relevant LCA activities as per the defined scope are included. The resource intensive steps such as future evolution of technologies are kept out of this first iteration. This is then followed by identification of influential aspects using global sensitivity analysis and which enables prioritization of improvements in data. Using the sensitivity indices and the newly introduced importance indices, most influential parameters and LCA processes are identified for improvement. Both qualitative or quantitative improvements can consequently be made in the models. Thus making the entire procedure more efficient.

In the following sub-sections, for each step of the methodology, the main actions to conduct are presented, based on the state of the art and on the potential uses of this model (i.e. design and optimisation in R&D).

Before proceeding further, it is important to have an overview of the LCA calculations. The impact (I) of energy conversion systems or of a single energy component, irrespective of the impact

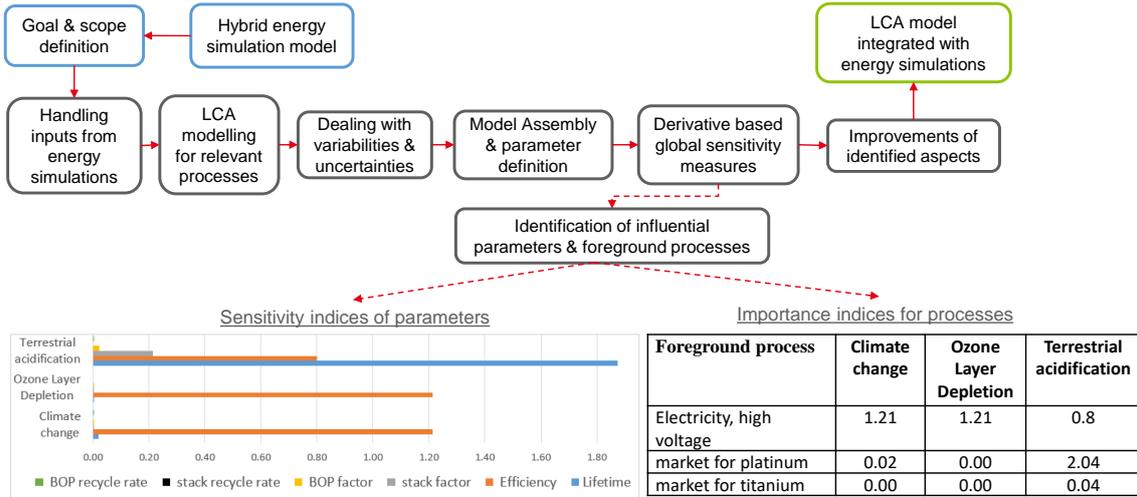


Figure 3.1: A schematic representation of the proposed framework. The steps that are described in this chapter are outlined in black.

category is calculated using the following equation:

$$I = \frac{\text{Impact over the entire life cycle}}{\text{Energy or fuel converted during the lifetime}} \quad (3.1)$$

Where the numerator is the sum of characterized impacts of each life cycle stage. It is evident that the LCA modelling calculates the environmental impact. However, the inputs to initiate these parameterized LCA models are calculated by the energy simulation. The denominator, on the other hand, is the part of the process flows calculated entirely by the technical simulation. The latter depends on local conditions, goal of the simulation and behavior of other components.

3.2.1 Handling inputs from energy simulations

The first step of the framework is to identify the outputs from the energy simulations and process them for use in LCA calculations. For a particular system configuration, once the goal and scope of the LCA investigation are decided, this step can be followed.

The components and their corresponding process flows that are to be included in the LCA calculations are identified. This has to be done according to the system boundaries defined in the goal & scope. Then, according to the components that are to be included, the outputs from the energy simulations are recognised. They have to be sufficiently processed before their utilisation in LCA calculations. So as to ensure that the units are consistent between the two.

At this point, additional inputs that are not included or calculated by the energy simulations but are relevant for LCA should also be added. Another important point is that there might be certain use phase modelling that affect the components/processes. It is crucial to take this into account. For example, the operation efficiency of a system might affect its lifetime.

3.2.2 LCA modelling approach

After the pertinent process and equipment are identified, LCA modelling approach for each them are decided in this step. From the point of view of LCA, energy simulations usually calculate the size of equipment or the quantities of flows related to the use phase. The upstream and downstream processes related to LCA generally require additional modelling. Three strategies are recommended in increasing order of specificity:

First is the direct use of activity from an LCA database or from another literature source. Here equivalence can be drawn with the ‘background processes’ as defined by ILCD handbook [European Commission2010]. From a management perspective, they are the ones where the designer does not have a direct influence or control on its selection. Or from a ‘specificity perspective’ when a specific process cannot be defined and average or equivalent data is to be used. For example, the designer might have direct control over amount of steel utilised in the system, but little information/choice where the manufacturer is sourcing steel from.

Second option is utilisation of an activity in literature with some changes. An average activity is modified according to the required goal and scope. Qualitative and/or quantitative changes can be made to the activity. There can be many applications to this due to multiple ways present to modify an activity. It is particularly relevant in this context for technologies whose flows are partially calculated by the energy simulations. For example, for crystalline silicon PV panels, energy simulations will calculate the electricity produced according to local weather data. Furthermore, other parameters such as lifetime or background electricity mix according to manufacturing data can be updated as well.

Third option is building a new LCA inventory altogether. This is particularly relevant for emerging technologies, whose inventories are not present in the LCA databases. In the next sub-section, recommendations to assemble inventory for this LCA are described.

Building an inventory for an emerging technology

The steps to assemble an inventory for an emerging technology in this context are different from a conventional LCA. Main reason being that this technology will only be a component when an integrated techno-economical-LCA analysis is conducted on the system. Keeping this in mind, this sub-section first establishes the background scope to collect the necessary data. Next, guidelines on the data search are provided.

Technology & Scope definition Technology definition as the name suggests is to define a technology whose model needs to be developed. This should be as specific as possible. For example, if the target is a model of PV panel, the type of panels, i.e. monocrystalline silicon, polycrystalline silicon or thin film should be specified. The following aspects have to be defined as precisely as possible:

Size The envisaged scale of technology accountable by the model has to be defined as well. For example, 50 kW – 2MW battery bank.

Geography The place where the system is installed and where the component was manufactured.

Temporality Since the context here is developing energy systems, the coverage here can be present scenario, near future or ex-ante analysis.

This section differs from ‘Goal & Scope definition’ of a conventional LCA by not specifying a goal and a functional unit. Instead a technology is specified instead. Especially since a model developed here will be a single component in the energy system. Furthermore, the size or geography might change during the optimisation or GSA analysis.

Finding inventory data

According to the anticipated technology readiness level (TRL) of components, guidelines on collecting inventory data are provided in this subsection. The methodology proposed needs at least one LCA existing in the literature for the energy component. It has been assumed that it is probable given the level of TRL in question (Demonstration or pilot scale).

The recommendation of this sub-section is to find enough data to build a complete inventory of the technology then refine the modelling in later steps. Here, complete means all (or most of) the materials and processes are present. The verification of completeness of the inventory can be done by one or a combination of the following options: data quality assessment in the original publications, multiple sources following similar data trends, expertise of engineers themselves or consulting an external expert in the technology domain. It ensures that no material or process is left out of the inventory at this step. The material amounts and their LCA modelling due to variabilities or uncertainties can be decided to be refined later.

To put completeness into context, the ILCD guidelines [European Commission2010] on LCA recommend a quantitative definition of ‘completeness’ of the inventory by defining a cut-off criteria for each environmental category either in the scope, or in an iterative manner during data collection/result analysis. In other words, declaration of the part of inventory left out of the analysis as a percentage (%) of the impact indicator. However, they also recognize the paradox, ‘if one would know the total impact, there would be no need for cut-off’ (ibid.). The important point is thus to ensure that the maximum activities are included in the analysis, to have appropriate results.

The search for inventory data can be challenging since there usually is a dearth of detailed and transparent data, even more so with emerging technologies. Multiple sources might be needed to complete one inventory.

The first step can be to verify if an article exists reviewing the LCA investigations performed on either of the following two topics: type of energy component or the function satisfied by the energy component. For example, for the case study of electrolysers, review articles on LCA of electrolysers or LCA of hydrogen production routes would be relevant. Apart from providing leads to the inventory data, these articles can provide important information required for parameterizing LCA models such as: major sources of variability in the impact and key issues to be resolved. In case where a review is not found, a manual search of data becomes necessary. However, it is suggested at this point to stop the search as soon as a complete inventory is built instead of finding all possible data sources in the literature.

Data quality and the relevance to the scope can be used to differentiate between multiple sources. Until then, average or anyone of the values found can be used. In the opposite case where no inventory is found, sources other than LCA studies have to be explored. There can be three main sources of obtaining them: scientific articles, reports from enterprises/organizations or obtaining data from laboratories or businesses [Arvidsson2018].

3.2.3 Dealing with variabilities and uncertainties

Once a preliminary inventory of the energy system is ready, there are different uncertainties including the inherent variabilities of technologies to be accounted for. In this subsection, variabilities of this context are anticipated and accordingly, best practices found in the literature are recommended. Regarding the types of uncertainties as per [Huijbregts1998] and [Björklund2002], the focus is on the following: spatial variability, variability between objects/sources, and uncertainty due to unrepresentative data.

Data gaps

Most of the time, even though we have a complete inventory, modelling these processes is not straightforward. It is possible that the processes to be added are not a part of LCA databases neither they are transparent enough in the literature to be replicated. Or, for some reason, detailed modeling is required to fill the data gap. For example, if a novel material has to be adapted for a future industrial production.

The steps below are recommended to be followed. They are in the increasing order of complexity. For instance, the fourth step includes the state of the art prospective LCA recommendations

on process design. Hence, they are advised to be followed when a particular inventory process is a major part of the technosphere or is identified as important later in this iterative framework.

First Chart the calculation steps in the original study.

Second Search the literature if another study has already modelled the particular process.

Third Otherwise, at this point, we can use an approximation using any of the following strategies: proxy or substitute processes, precursors, by-products, co-products, ideal baseline (stoichiometric ratio and heat of reaction for mass and energy balance respectively) or expert opinion. LCA practitioners use these techniques widely.

Fourth After the previous step, it is expected, that the data gap is filled. Even though questions could be posed about accuracy, it can be refined iteratively. Hence, after the DGSM step, if a foreground process is recognized as important, a more detailed estimation is required. Here if other strategies are not feasible, following approaches could be useful:

- Process design approaches in LCA can be practical in this case. They are also the most relevant to the type of processes we intend to model (manufacturing, recycling, etc.). They use ideal baseline, publications or lab scale data as starting points. From the literature, there are a variety of different approaches proposed when the data is lacking [Hischier2005, Simon2016, van Kalkeren2013, Hetherington2014, Tecchio2016]. Most relevant approach found was of [Piccinno2016], where guidelines for calculation inventory data based on process design are included. It consists of a simple scale-up framework based on laboratory or ideal baseline data to provide quantitative estimates. Calculations for commonly used components in manufacturing along with average and recommended values to facilitate heat/energy balance are also included (ibid.).
- Process design using simulation tools can be also be used instead or as a complementary approach. An example could be process simulation software such as Aspen ¹ or ProSim ², where the input/outputs of a process could be estimated using energy and mass balance. Another useful but not necessarily a process design tool could be Finechem ³, available freely to estimate impact of petrochemicals using molecular descriptors. However, this particular tool should only be used for data with low importance [Wernet2012].
- For the impacts of metals not in the LCA databases, either the values or inventory modelling process from Nuss and Eckelman could be used [Nuss2014]. They present cradle-to-gate impacts of 63 metals, by modelling inventories using ecoinvent version 2.2 and the data collected from literature.
- To estimate 'fugitive' emissions'. When modelling processes concerning volatile compounds, some amount of them 'escape' or 'leak' into the atmosphere. They are called fugitive emissions and can have a significant impact on the environment [Thomassen2019]. To model them, Hassim and colleagues propose emission factors that can be applied to a process design [Hassim2010]. Hence, they should be used where relevant.

A detailed, step-by-step guidance on finding inventories and assembling them can significantly support engineers, who are beginners in LCA. Such a guideline, does not exist in the literature so far.

¹ <https://www.aspentech.com/en/products/engineering/aspens-plus>

² <https://www.prosim.net>

³ <https://emeritus.setg.ethz.ch/research/downloads/software—tools/fine-chem.html>

Scaling

The recommended procedure for scaling inventories is elaborated in this section and justified as much as possible. An illustration can be seen in figure 3.2, the procedure is adapted from the one proposed by [Gerber2011], modified according to the context. It is described as follows. The ideal case naturally is to obtain direct material distributions for different sizes, but finding such data is not expected in this case. Hence, we start from the next best step, estimating material amounts using design or engineering laws.

If this is not possible either, the scaling equation 3.2 can be used with the subsequent steps. First step is to find the most relevant parameter for scaling the inventory activities. Often it is the sizing parameter, which the designers are expected to be familiar with. For example, energy generation components, scale with the nominal power (kW or MW), while for storage, it is energy storage capacity (kWh or MWh).

$$\frac{L_{y,co}^{a,2}}{L_{y,co}^{a,1}} = \left(\frac{SP_2}{SP_1} \right)^b \quad (3.2)$$

Where $L_{y,co}^a$, is the quantity of activity 'a' in the inventory of component 'co' in its lifecycle phase 'y' and SP is the sizing parameter of equipment (e.g. kW or kWh), b is the scaling factor.

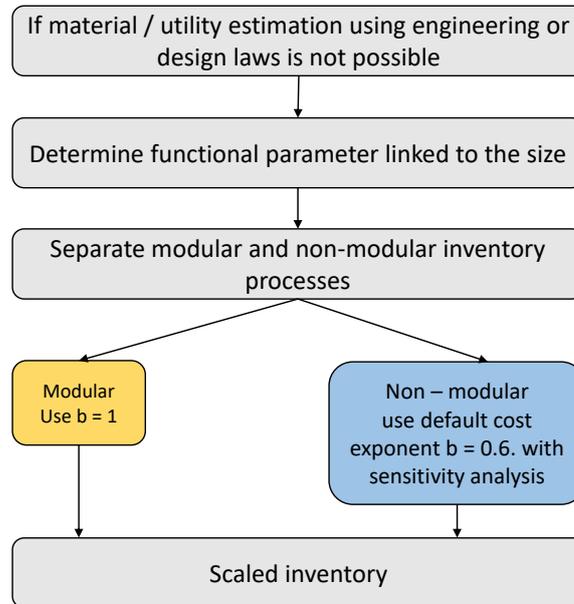


Figure 3.2: Proposed scaling approach for dimensioning inventories of components. Adapted from [Gerber2011].

Then, a classification of inventory activities is made under modular and non-modular components. Modular components are the one where a set of structures are repeated to increase the capacity. As result, the 'economies of scale' don't play a role and amount of material per unit size of component follows a linear correlation (e.g. electrolyser stack, fuel cell stack, solar PV panels, etc.). Hence the inventory activities under these will scale with $b=1$. Non-modular components are the opposite, where there is marginal decrease/increase in materials per unit size (e.g. heat exchangers, boilers, compressors, wind turbines). For non-modular type, cost exponents can be used. If these are unavailable as well, the default cost factor of $b = 0.6$ can be used. However, in this case a sensitivity analysis assessing the influence of inventory uncertainty becomes essential, in order to assess the robustness.

It should be noted that a single technology could have both modular and non-modular components. The below equation is sometimes also called '6/10th rule' when $b = 0.6$ is used. It is widely

used for cost scaling if no other information is available [Moore1959, Chauvel2003].

The above recommendations are based on the literature where few studies analyse the influence of size of components on the environmental impacts. Caduff *et al.* assessed different energy conversion units such as biomass furnaces and heat pumps [Caduff2014]. They found a clear correlation between the impacts (global warming potential and ozone depletion) during the manufacturing phase of an equipment and its size, according to power law in equation 3.2.

Similar results were obtained, with different 'b' values was found for mass (kg) of energy components such as diesel engines, generators and steam boilers, including the ones mentioned above [Caduff2011]. These values lie in the same range as the recommended cost scaling factor 0.5-1 [Caduff2014].

Gerber *et al.* found a similar equivalence between cost and impact scaling of equipment [Gerber2011], further justifying the use of equation 3.2. They also say that cost exponent although a reasonable estimation, is still an approximation, which brings additional uncertainty. This uncertainty can be due to cost dependence on factors such as labour. For the value of b, Gerber *et al.*, recommend the use of empirical data calculated using multiple inventories or using cost-scaling factor. Hence, use of engineering or theoretical design laws for material estimation is the recommended before.

The difference between the strategy presented here and that of Gerber *et al.* and Caduff *et al.* is the manner of using the power law in equation 3.2. Here, the individual activities are scaled after their classification as modular/non-modular. Whereas, in Gerber *et al.*, aggregated impacts of components are scaled directly with respect to the size. In other words, in equation 3.2, impact scores of components (e.g. global warming potential) are on the left hand side, instead of individual inventory amounts. Caduff *et al.* demonstrate multiple options which could be used without specifying one.

Another important point is that different indicators may scale differently. That is to say, with a different scaling factor 'b'. Dissimilar scaling factors of Caduff *et al.* [Caduff2011, Caduff2014] obtained for global warming potential, ozone depletion and mass of components, are an evidence of this. The proposed approach provides more flexibility, as it enables using power law for some processes and the possibility of using other approaches for the rest. It is also more logical as we will see in the example of electrolyzers, where the stack and auxiliary components scale differently.

Hence, we put forward a more flexible and normative approach contributing to the scaling of energy components for LCA calculations.

Geography

The ideal scenario to adapt the inventory activities would be possible if the supply chain of the components is known. In general, at least in ecoinvent v3.6 [Wernet2016] it was found that there were a lot more choices for varying geographic scope for activities related to energy than materials. The recommendation is thus to select inventory activities as close to expected scope as possible. They can be further be customised/refined in the databases if they are found to be influential later. Processes such as the ones below can be chosen/modified according to the geography using the LCA databases:

- Electricity mix
- Metals, chemicals and other materials
- End of life scenarios

In hybrid energy simulations, the variability due to meteorological data is already taken into account. The geographical scope can have a significant impact on the LCA results. For example, the influence of electricity mix utilised in PV manufacture varied the results of carbon footprint from approx. 40 g CO₂-eq/kWh for EU-hydropower mix to 80 g CO₂-eq/kWh for Chinese mix, keeping other parameters the same [Wild-Scholten2013].

Future evolution

Depending on the case, we might or might not need to adapt our model for the future. Subsequently, users can look into opportunities at adapting them for the future. Following are the aspects that could be altered:

- Improvements in manufacturing and process efficiencies – As the cumulative deployment of a technology increases, the cost of technology has been observed to decrease. This ‘learning’ is because of various reasons such as: increased productivity, industrial scale production, research, etc. These factors consequently also affect the environmental impacts. They can be modelled using the ‘one-factor learning curve’ [Rubin2015] :

$$I_{n,Cum} = I_{n,0} \cdot Cum^c \quad (3.3)$$

Where, $I_{n,0}$ is the impact for the indicator ‘ n ’ (e.g. kg CO₂-eq/kW) of the first component. While $I_{n,Cum}$ is the impact of the last component at a cumulative installed capacity Cum (e.g. kW) respectively. Technological learning rate is ‘ c ’.

In life cycle assessment, this equation has been applied to project impacts of wind power [Caduff2012] and PV panels [Bergesen2016, van der Hulst2020]. However, they are not advised to be used when the learning rate ‘ c ’ is calculated with low-quality data. Where, low-quality defined by Thomassen *et. al* as <3 data points or low co-efficient of determination ($R^2 < 0.5$) [Thomassen2020]. Mainly since the predictions are uncertain and highly sensitive to the learning rate ‘ c ’ [Yeh2012]. Furthermore, unlike economic models, there are no sector-wise rules of thumb for the environmental impact [Thomassen2020]. Hence, they have limited utility for emerging technologies.

Thus the advise in this case would be consulting experts/stakeholders for expected material and process efficiency gains. Another possible approach could be to estimate the improvements based on a similar mature technology. Such as reductions expected in stack materials for PEM electrolyser based on PEM fuel cells [Carmo2013].

- Evolution of background activities – The background activities such as the electricity mix or end of life scenarios can be adjusted for the future. Moreover, if they lead to large variance in the results, comparison using different ‘predictions’ is advised [Arvidsson2018]. Some examples of the documents can be roadmaps, goals, prediction scenarios of IEA, scenarios of environmental agencies, scenarios of NGO’s, etc.
- Energy conversion efficiencies – These involve the use phase efficiency of components, for example, efficiency of solar panels. The engineers are already involved in techno-economic estimations where these values are necessary. In any case, roadmap/prediction documents or expert opinion can be consulted.

To conclude, all of the above approaches add another layer of uncertainty and require significant amount of effort from the practitioner. Hence, for the first iteration, it is advised to use the easily available data. Once sensitivity indices are available, understanding of the model becomes clear and then we can focus on detailing the modelling that significantly influences on the results.

3.2.4 Model assembly and parameter definition

At this stage, it is expected that the outputs from the energy simulations are appropriately transformed for LCA calculations. The LCA modelling is finalised for the first iteration that includes all the processes under the system boundaries.

According to the identified uncertainties and variabilities in the previous subsection, the LCA model is parameterized i.e. defined in terms of parameters. There can be multiple parameters possible for a model. Some parameters are already available with the energy simulations. Generally, these are the ones related to the operational phase of the system. The accuracy or the quality of energy systems modelling will thus invariably affect the LCA results. Thus, care has to be taken to include all the flows influencing the LCA impacts.

In addition to them, parameters related to LCA might also have to be added. They could be a design parameter to be fixed or could also be processes lying outside the scope of influence of designers. A typical example can be electricity mix, where the designer has little influence over its development, however, it may play a crucial role on the impact of the system in the future. These parameters can be any of the variabilities or uncertainties in the LCA model.

The computational structure of parameterised LCA models can be defined in multiple ways, which were presented in section 2.2.4. Irrespective of the chosen manner, two crucial points are highlighted here regarding the model assembly which will be essential for the next steps of sensitivity analysis:

1. It should enable the computation of contribution analysis in terms of each life cycle phase of each component.
2. There should be at least one parameter used for sensitivity analysis in DGSM for the first iteration that is directly linked to the contribution of one component to a particular life cycle phase. The parameter range can be decided using real world data or from the literature.

To elaborate further, consider figure 3.3 where the equation 3.1 can be viewed in terms of its contribution from each life cycle phase. This contribution should be further separated in terms of the individual components in each phase if possible. For example, if the manufacturing phase consists of two separate components: boiler and gasifier, the contribution analysis should be made possible for manufacturing phase of each the two, with at least one parameter related to each one. For the implementation of this framework, such a representation is important. This will help in the next section to identify the influential LCA processes.

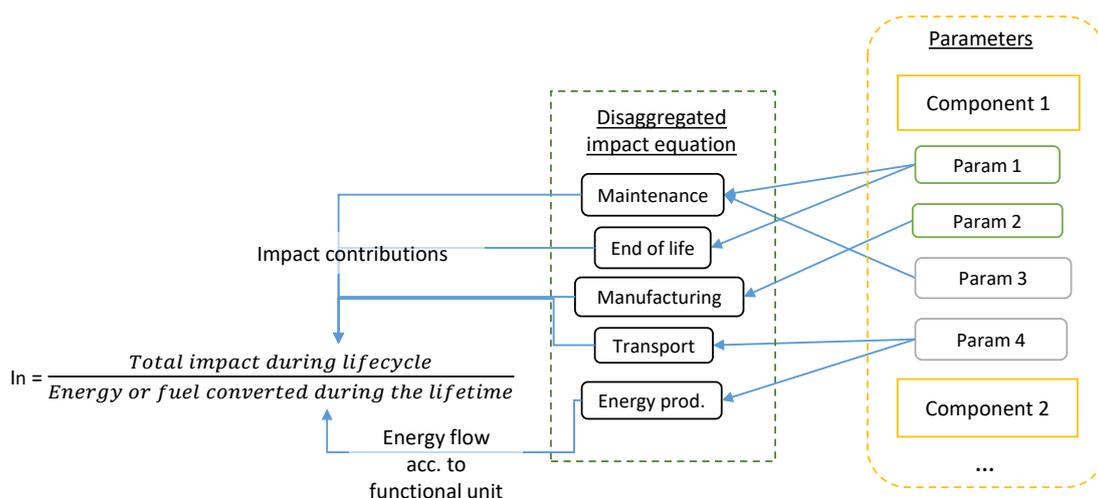


Figure 3.3: Representing the parameterized model and disaggregating the impact equation in terms of impact contributions and energy flow.

A caution to the use is the specification of range of uncertain parameters. If an unrealistic large range is specified, the method can assign higher influence to the parameter. In the opposite

case, when shorter range than actual is considered, lower influence might be obtained than that is actually possible. Thus the uncertainty ranges should be based on realistic ranges from literature, real world conditions or expected values.

3.2.5 GSA using Derivative-based Global Sensitivity Measure

In this step, the parameters selected along with their uncertainty ranges are supplied to DGSM method. Then, the influential parameters and processes are identified using the obtained sensitivity indices and importance indices respectively. These 2 aspects are then improved according to required scope. More than one iterations might be required depending on the case. If in case, the value of an identified influential parameter cannot be fixed, the uncertainty values can be propagated to the model. Resulting in a distribution, which can be analyzed using statistical values. This will enable users predict the possible impact range.

In this section, first the choice of DGSM as a global sensitivity analysis method is justified. This is especially relevant since it has not been used yet in the LCA literature. Then, the mathematical description and practical information about its implementation are discussed. Finally, the importance indices are introduced.

Preference of using GSA over SA

Sensitivity analysis (SA) is recommended for all LCA's. Due to the added layer of uncertainties thanks to the ex-ante context, SA becomes crucial. If relative influence of parameters is identified, it allows prioritisation of resource allocation. Moreover, it enables users to judge the accuracy of their models. In summary, we utilize it for the following reasons:

- Understanding of the model behavior according to its influential parameters
- Prioritizing data collection
- Enhancing robustness of results when design parameters are uncertain

Even though the most common SA in LCA's is the local sensitivity analysis (LSA), where one input is varied around its reference point, keeping other values the same [Groen2017]. Their interpretation is limited since they assume model linearity, ignore interactions between parameters and do not search the entire distribution space [Pannier2018]. Even though it can be easily applied, it cannot quantify the responsibility of uncertain parameters in the model output variability [Nadal2020]. Global sensitivity analysis (GSA) methods do not have these drawbacks [Wei2015]. These methods evaluate the model for different combinations of uncertain parameters, unlike the local SA. Furthermore, uncertain parameters not only from LCA modelling but also from energy simulations are expected. Thus, in a complex model such as this, GSA makes much more sense.

Selection of DGSM as a GSA method

Other GSA can also be used, but in this paper, we recommend the use of Derivative-based Global Sensitivity Measure (DGSM) [Sobol'2010]. The logic behind this choice is that the goal is to discard the non-influential parameters, and obtain their approximate ranking. It offers a good balance in this context between accuracy in calculating sensitivity indices and computational time. This choice is further elaborated with respect to other GSA methods as follows.

In this context, an approximate ranking of the influential parameters and discard of non-influential ones is needed, which can be done while maintaining a low computation time using DGSM [Lamboni2013]. This is especially important in this context for two reasons. First, the hybrid energy simulations are computationally expensive due to dynamic simulations. Due to the

variable nature of renewable energies and variable demand, an energy balance at each time-step becomes necessary. Second, in this context, due to the large amount of uncertainties and variabilities, number of uncertain parameters can be 10+.

To put the time required in context, for 10 parameters (d), no. of model evaluations for variance decomposition method of Sobol are $N(2d+2) = 11264$, with $N = 512$ [Saltelli2008]. It is widely considered as the state-of-the-art in GSA. While for DGSM, less number of model evaluations are needed with $N = 100$, $N(d+1)=1100$. If one model evaluation takes around 15 seconds, thus the time required for Sobol is around 47 hours, while for DGSM, 4.6 hours. Apart from the ones mentioned, other GSA methods exists as well, but DGSM can work with a black-box models, and offers the best trade-off between required accuracy and acceptable computational time in this context. Another advantage, unlike regression/correlation based methods, is that it can be used with non-linear or non-monotonic models (ibid.).

DGSM can be seen as a generalized form of the popular screening Morris method, also known as, method of elementary effects (ibid.). In other words, Morris method is a coarse form of DGSM. Morris method may require lesser number of model evaluations than DGSM [Iooss2017], but in its basic form, it cannot precisely compute the sensitivity due to interaction between parameters, neither can it rank them in order of their influence [Lamboni2013].

Moreover, a link between DGSM and the method of Sobol has been established [Sobol'2009]. Meaning that even in the worst case it allows reliable discard of non-influential parameters and even rank parameters in the order of their influence (except for highly non-linear models).

Thus even in the worst case, the non-influential parameters can be reliably discarded but the ranking of some parameters is inversed. Its not a problem since exact knowledge of variance decomposition is not necessary in this case.

Mathematical description

Let's say, $f(\mathbf{P})$ is the combined techno-economic-LCA function to be evaluated with d parameters, where $\mathbf{P} = (p_1, p_2, \dots, p_d)$ and $e = 1, 2, \dots, d$ Thus \mathbf{P} represents a set of parameters. As per [Sobol'2009], DGSM measures are mathematically defined as:

$$v_e = \int \left(\frac{\partial f(\mathbf{P})}{\partial p_e} \right)^2 d\mathbf{p} \quad (3.4)$$

Since, we use a black-box model, the right hand side of the above equation can be numerically calculated as per [Iooss2017, Touzani2014, Lamboni2013]:

$$\hat{v}_e = \frac{1}{N} \sum_{j=1}^N \left(\frac{\partial f(\mathbf{P}_j)}{\partial p_e} \right)^2 \quad (3.5)$$

$$\frac{\partial f(\mathbf{P})}{\partial p_e} = \frac{f(p_1, \dots, (p_e + \delta), \dots, p_d) - f(\mathbf{P})}{\delta} \quad (3.6)$$

In simple terms, the function $f(\mathbf{P})$ is assessed with small increments ' δ ' in a parameter p_e in equation 3.6. Then, v_e becomes the mean over N repetitions of different modifications p_e for the entire parameter distribution using equation 3.5. Since we are using the python package SALib [Herman2017], this distribution of parameters are developed using Sobol sequence over a $N(d+1)$. Other sequences to generate samples can also be used instead [Kucherenko2017].

In this thesis, we use γ_e as the sensitivity index because of the following link established with the Sobol total effect, $S_{e,total}$ [Sobol'2009]. The total effect index of Sobol, considers the non-linear and effect of interactions between parameters. A uniform distribution is used for all parameters, where a_e and b_e are the lower and upper bounds of p_e . Whereas, V is the total variance

of the distribution. Another probability distribution can also be used, but finding a distribution or large enough dataset for its calculation is not anticipated due to the scarce data.

$$\gamma_e = \frac{(b_e - a_e)^2}{V\pi^2} v_e \geq S_{e,total} \quad (3.7)$$

Using the above equation, we see that for low values of γ_e , the parameter can be considered non-influential. Thus, screening of parameters, while taking a fraction of time as compared to the calculation of the sobol indices $S_{e,total}$. The conclusion being that even in the worst case of highly non-linear models (e.g. x^5), low values of sensitivity indices γ_e can reliably confirm non-influential parameters.

Practical implementation

According to a recent review [Douglas-Smith2020], DGSM is readily available in sensitivity packages in the software R [Iooss2019] and Python (SALib) [Herman2017]. In this paper we use the latter. The utilisation is simple, once the parameterized model is finalised, it can be linked to sensitivity analysis package. The number of uncertain parameters and their ranges have to be specified. First, a distribution of various combination of these parameters is obtained. Then, samples are generated by propagating these parameter combinations through the model. The package treats the integrated technical-economic-LCA model as a black-box function. Finally, the generated samples are mathematically assessed as per the previous section to get the sensitivity indices. The process is outlined in figure 3.4 Number of model evaluations required are $N(d+1)$, where d is the number of parameters. While $N=100$ can be used to get consistent results as per investigations done in the past [Touzani2014, Iooss2012, Iooss2017].

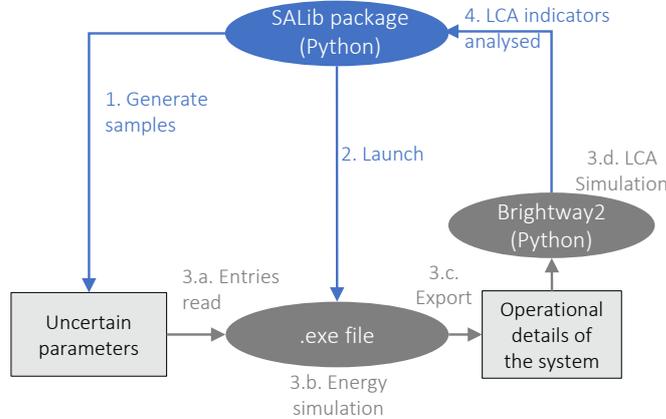


Figure 3.4: Interaction of sensitivity analysis package with energy simulations and LCA modelling.

Identification of environmentally influential parameters and foreground processes

Here we focus solely on LCA indicators, even though other indicators such as levelised costs are also possible. Influential parameters are identified using sensitivity indices while a novel coefficient, importance index (Im) is proposed to identify the most influential foreground processes. The sensitivity index, γ_e obtained from DGSM will enable us to discard non-influential parameters. Thus, utilizing resources on improving them and their related term in the fraction has a lower priority. On the contrary, the influential parameters have to be estimated accurately to obtain appropriate results.

Apart from the parameters, foreground process modelling might also require improvements. To facilitate this exercise, importance index (Im) of an activity 'a' can be defined as follows for a particular LCA impact indicator:

$$Im_{y,co}^a = \gamma_e \times C_{y,co}^a \quad (3.8)$$

Where, γ_e is the sensitivity index of the parameter p_e which is linked to a component or process (co) in the lifecycle phase y . While $C_{y,co}$ (%) is the contribution of the process a to the total impact of component (co) in lifecycle phase y . This is possible since each of the selected variable is linked to a life cycle phase.

For example, consider a hypothetical component, bio-PV panel. Let's say, a design parameter, panel thickness was varied, to ease visualisation, consider it as LCA parameter 2 in figure 3.3. Suppose it has the highest sensitivity index, $\gamma_{manf, thick} = 0.80$. Meaning, it is responsible for majority of the variance in results. It is directly related to panel manufacturing where say, natural gas use contributes to 90% of the global warming impact. If the material proportions are unchanged during the variation of thickness, $Im_{bio-PV, manf.}^{nat.gas} = 0.8 * 0.9 = 0.72$ for the category global warming potential.

Here we establish for the hypothetical case that, modelling of natural gas is important in this case to improve model accuracy. Hence, user could improve the estimation of natural gas; assign a better representative natural gas process; or even change the composition of natural gas process in the future. In case one process is related to more than one influential parameter, the index is then the aggregated sum. Importance indices for each indicator for the foreground processes can thus be presented in decreasing order. Value of sensitivity and importance indices close to zero represents they are non-influential.

Importance indices can also be seen as a new method to quantify the importance of uncertainties in LCA. They can be especially useful when large amounts of uncertainties are present in the system. The foreground processes whose improvements will have the maximum influence on the results can thus be identified. Higher the co-efficient, the more resources should be allocated to improve the particular process modelling. They are not sensitivity indices in themselves, but they can provide useful information as to which processes should be prioritized for modelling.

Thus, the influential parameters and LCA activities are identified using sensitivity indices (γ) and importance indices (Im). They can thus be differentiated in terms of their relative importance to the final results. Data can thus be preferentially searched and the resources can be better allocated for LCA modelling.

3.3 Case study: PEM water electrolysis

For demonstration, the framework is applied to the same case study of hydrogen production from PEM electrolysis, which was introduced in section 1.1.5. An illustration of LCA boundaries in the case study and in the Odyssey software is presented on left and right respectively in figure 3.5. The cradle-to-grave impacts of electrolyser and power supply is considered within the LCA boundaries. Hybrid energy simulation model is used as an input to the parameterized LCA. Any other simulation or software could also be used instead. While the LCA calculations are done using Brightway2, an open-source LCA framework in Python [Mutel2017]. Python package of SALib was used for DGSM analysis [Herman2017].

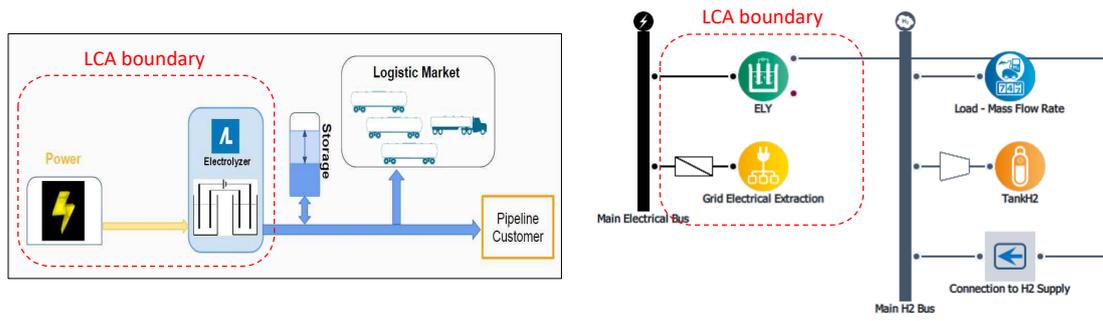


Figure 3.5: Illustration of the case study and the LCA system boundaries

3.3.1 Handling outputs from the energy simulation

For this case study, two components to be included for LCA calculations can be explicitly found looking at the figure above: PEM electrolyser and grid electricity. On closer look, two process flows from the electrolyser are also found: water consumption and hydrogen production. The latter is the output and impact of all the components has to be calculated with reference to it. Thus there are three components/processes whose cradle-to-grave impact has to be calculated.

The relevant outputs from the energy simulation models are time series: Stack power (E_t^{PEM}), Hydrogen production ($Prod_t^{H_2}$) and electricity extraction from grid (E_t^{grid}). Where t is the time step, it represents data at every 5 minutes for one year. Time steps can be denoted by $t \in T = \{0, 1, 2, \dots\}$, with its length thus becoming $len(T) = 105120$.

They are used to then used to calculate the lifetime consumption/production values of Grid electricity consumed during the entire life time (E_{life}^{grid}), and hydrogen production ($Prod_{life}^{H_2}$) and size of PEM electrolyser (S_{PEM}) using the following equations:

$$S_{PEM} = \max(\{E_t^{PEM}, t \in T\}) \quad (3.9)$$

$$E_{life}^{grid} = \frac{len(T)}{12} * lifetime^{project} * \text{mean}(\{E_t^{grid}, t \in T\}) \quad (3.10)$$

$$Prod_{life}^{H_2} = \frac{len(T)}{12} * lifetime^{project} * \text{mean}(\{Prod_t^{H_2}, t \in T\}) \quad (3.11)$$

Where, $lifetime^{project} = 20$ years. Water consumption over the entire lifetime is calculated using the hydrogen production. Since 1 mol of water (mol. wt. = 18g) produces 1 mol of hydrogen (mol. wt. = 2g), nine times more water is consumed than hydrogen produced:

$$W_{life} = 9 * Prod_{life}^{H_2} \quad (3.12)$$

Then, it was found during the preliminary analysis that the PEM electrolyser stack might have to be renewed during the project lifetime of 20 years. This depends on the number hours of operation. For this, first the Number of hours ($hours^{PEM}$) of stack functioning are calculated using the number of time steps when the stack power is greater than zero:

$$hours^{PEM} = \frac{len(\{E_t^{PEM} > 0 | E_t^{PEM}, t \in T\})}{12} * lifetime^{project} \quad (3.13)$$

Then, the number of stacks required during the project lifetime are estimated using the equation below, where ceiling means the integer greater than or equal to the calculated real number. For example, if 2.3 stacks are calculated, that actually means 3 stacks will be required during the lifetime of the project. $Stack_{life}$ is the stack lifetime in hours.

$$Number_{stacks} = ceiling \left[\frac{hours^{PEM}}{Stack_{life}} \right] \quad (3.14)$$

3.3.2 LCA modelling approach

Then for the LCA modelling approach, out of the three whose impact needs to be calculated, two activities, grid electricity and de-ionised water are interfaced directly with the simulations for the first iteration. They are classified under the use phase activity list as L_{use}^a , where a refers to a particular activity in the list. For PEM electrolysis, no activity is found in the ecoinvent v3.6 database. Hence, an inventory is built with the following steps:

Building inventory for PEM electrolyser

Technology and scope definition With multiple demonstration systems coming up around the world, TRL level of PEM electrolyser can be considered to be 7-8. Further, we select the following scope:

- Scale of electrolyser model: 100 kW - 10 MW
- System is installed in Marseille, France and it is assumed to be manufactured in the same country
- The system is to be installed in the year 2030 with a project lifetime of 20 years

Search for inventory data Following the first step in this subsection, multiple review articles about hydrogen production were found [Valente2017, Bhandari2014]. In these papers, only one study [Schmidt Rivera2018] was found which could be used to partially reproduce the inventory. The study enlists data of the PEM stack but does not contain any information about the auxiliary or BOP (balance of plant) components. These components also referred to as balance of plant, are essential for hydrogen production. They mainly involve electrical components, pre-processing of water and post processing of hydrogen [Pilenga2018].

Hence continuing the manual search, two reproducible inventories comprising of all the expected components were found [Bareiß2019, Zhao2018]. The completeness of the inventory was verified by an expert, Dr. Fouda-Onana. Being involved in the R&D activities of PEM electrolyser at CEA-Liten, he is one of the contributors to harmonization methodology of low-temperature water electrolysis, proposed by European Commission [Pilenga2018]. The exchange was held in the form of in-person meeting, with the primary goal: verifying if the inventory is consistent with the components considered in the system boundaries. Other aspects were discussed such as: Nafion manufacturing process, current efficiencies and lifetime of PEM systems and the expected future developments in the inventory. At the end of the meeting, the (qualitative and quantitative) completeness of the inventory was verified and publications were shared to improve the understanding/modelling of PEM electrolysis systems. Thus the PEM stack and BOP inventory can be found in annex tables A.1 and A.2 respectively. The amount of each activity can be represented by $L_{manf, stack}^a$ and $L_{manf, BOP}^a$, they take into account processes of raw material extraction, manufacturing and assembly.

Even though the stack data was available in [Schmidt Rivera2018] it is not used since the other two studies align better with the scope while having a higher data quality. First for the goal and scope, the two studies [Bareiß2019, Zhao2018] have a goal of industrial hydrogen production with 1MW PEM electrolyser. On the other hand, for Schmidt-rivera *et al.*, electrolyser size can be estimated to be < 100 kW, for domestic use. In terms of data quality, their stack inventory is based on a lab report from NREL in 2005 [Carlson2005], while the selected two studies estimate inventory from pilot PEM systems based on relatively recent data, personnel interviews and industrial visits.

3.3.3 Addressing variabilities and uncertainties

Data Gaps Once the foreground model was assembled, there were issues with three activities: Nafion membrane of electrolyser, deionizer and iridium. Since they were not available in ecoinvent v3.6 (latest at the time). Hence, they were dealt in the following approach for the first iteration:

- Nafion is modelled according to the stoichiometric calculations of [Duclos2017]. The membrane is considered to be a product of sulfuric acid and tetrafluoroethylene with 100% yield and negligible energy requirements.
- Deionizer – Instead of modelling a deionizer separately, the activity ‘deionised water’ was used.
- Iridium is a part of platinum group elements and is mined along with other rare metals of this group: platinum, ruthenium, osmium, palladium and rhodium. Hence, ‘platinum’ is considered a substitute for the impacts of iridium.

As we can see, the modelling is done with the easily available data for the first iteration.

Scaling Engineering models to estimate the amount of materials in the PEM stack and BOP were unavailable, hence the equation 3.2 is used with a factor $b=1$ for stack inventory and $b=0.6$ for BOP. Since the stack power is increased modularly, while there is decrease in material quantity per unit capacity for BOP components. Other studies use similar hypothesis for electrolysers [Zhang2017, Mehmeti2018].

Geography and End-of-life (EOL) For the geographical scope, all the inventory materials, were selected from the ecoinvent 3.6 cut-off database, with location priorities defined as follows: France, neighbouring countries, Europe and finally global markets. It should be noted that the activities for stack and BOP are based on the market mixes from ecoinvent i.e. they include materials from various sources, transport, etc.

At the end of life, the materials in the PEM stack and BOP are either recycled or disposed (landfill or incineration). The system is credited for the primary production avoided due to recycling. In order to find activities for these steps, the recycling (secondary production) $L_{EOL, recycle}$ and primary production activities $L_{EOL, primary}$ for each material in inventory is gathered from ecoinvent. Furthermore, the disposal activities $L_{EOL, dispose}$ are also assembled where the waste markets specific to each material were selected. Since in the selected waste markets, according to the material, they were already classified as incinerated or landfilled or a mix of both. If waste market was not found, the material was assumed to be landfilled in a generic activity. The entire list of the three types activities can be found in annex tables A.3, A.4 and A.5.

End of life impact of stack and BOP materials is calculated separately as a function of the found activities and the recycle rate. The following steps are thus followed:

- Each material a in the electrolyser manufacturing inventories $L_{manf, stack}^a$ and $L_{manf, BOP}^a$ is iterated through the recycle list $L_{EOL, recycle}$.
- If a relevant recycle process for a exists in $L_{EOL, recycle}$, then it is added to the recycle list of that component with the quantity calculated as:

$$L_{EOL, recycle, co}^a = R * L_{manf, co}^a \quad (3.15)$$

Where R is the recycle rate. Same amount is added to the avoided list of that component but as a negative input for LCA calculations as:

$$L_{EOL, avoid, co}^a = -R * L_{manf, co}^a \quad (3.16)$$

Rest of the material is disposed using a relevant process in $L^{dispose}$ as:

$$L_{EOL\ dispose,co}^a = (1 - R) * L_{manf, co}^a \quad (3.17)$$

- If a relevant recycling process is not found, entire material quantity is disposed using a particular disposal activity if possible or a generic landfill activity.

$$L_{EOL\ dispose,co}^a = L_{manf, co}^a \quad (3.18)$$

The amount of recycled materials incorporated the manufacturing of metals in France was around 50%, while it was only 6% for plastics [ADEME2018]. This does not represent the amount of materials actually recycled, but it is taken as an approximation for the recycle rate R for the first iteration. This can also be improved later if required. Finally, three EOL inventories each for PEM stack ($L_{stack}^{recycle}$, L_{stack}^{avoid} , $L_{stack}^{dispose}$) and PEM BOP ($L_{BOP}^{recycle}$, L_{BOP}^{avoid} , $L_{BOP}^{dispose}$) are formulated using the equations presented above. End of life impact of each material thus equals impact of recycling plus impact of disposal minus the avoided production of primary material due to recycling.

Future evolution As discussed in the methodology, future evolution of the LCA processes is intentionally left out of the modelling for the first iteration. Even though some information is available in the references found thus far. Further discussion in section 3.4.2.

3.3.4 Model Assembly

Finally, the life cycle impact has to be calculated for the activities along all the life cycle phases. This can be represented as a demand array \mathbf{f} for the LCA computations:

$$\mathbf{f} = (L_{use}^a, L_{manf, BOP}^a, L_{manf, stack}^a, L_{EOL\ recycle, BOP}^a, L_{EOL\ avoid, BOP}^a, L_{EOL\ dispose, BOP}^a, L_{EOL\ recycle, stack}^a, L_{EOL\ avoid, stack}^a, L_{EOL\ dispose, stack}^a) \quad (3.19)$$

Equation 2.4 is recalled, which is utilised for LCA calculations using Brightway2. The total life cycle impact of the system is thus sum of the impacts of the above list of activities for a particular impact indicator. The impact per unit hydrogen is found by dividing the total hydrogen produced, i.e. equation 3.1. It should be noted that other ways of parameterised LCA computations discussed in section 2.2.4 could also be utilised which will provide the same results.

Five parameters are made directly related to at least one of the lists so that the relationships between them and the contribution analysis could be made explicit. It will later be useful for the sensitivity analysis step. Only one parameter, efficiency (in grey), was already present in the energy simulations. Four other parameters relevant to LCA are added in green: stack lifetime, recycle rate and material scaling models of stack and BOP. A simplified version of global structure of the parameterized model is illustrated in figure 3.6. The disaggregated impact equation can be seen consisting of contributions of different components/phases divided by the hydrogen production during the lifetime. The relationships between the disaggregated impact equation and the parameters can also be seen clearly. Other parameters also exist in energy simulations but only the ones that are varied in this investigations are illustrated.

Recently updated impact assessment method, IMPACTWorld+ [Bulle2019] is used. The choice is based on the fact that, this impact assessment framework was the latest to be updated according to the best practices in the characterization modelling. However, other impact assessment methods could alternatively be used. Hydrogen is produced from electrolysis primarily to avoid using the fossil sources. Hence, we select climate change indicator to quantify this potential. Furthermore, since PEM electrolyser consists of many rare metals, following midpoint indicators are also selected: terrestrial acidification, particulate matter formation, photochemical oxidant formation, human toxicity cancer and mineral resources use. Ozone depletion could also be important since the PEM stack consists of Nafion membrane, which is manufactured using fluoro- compounds.

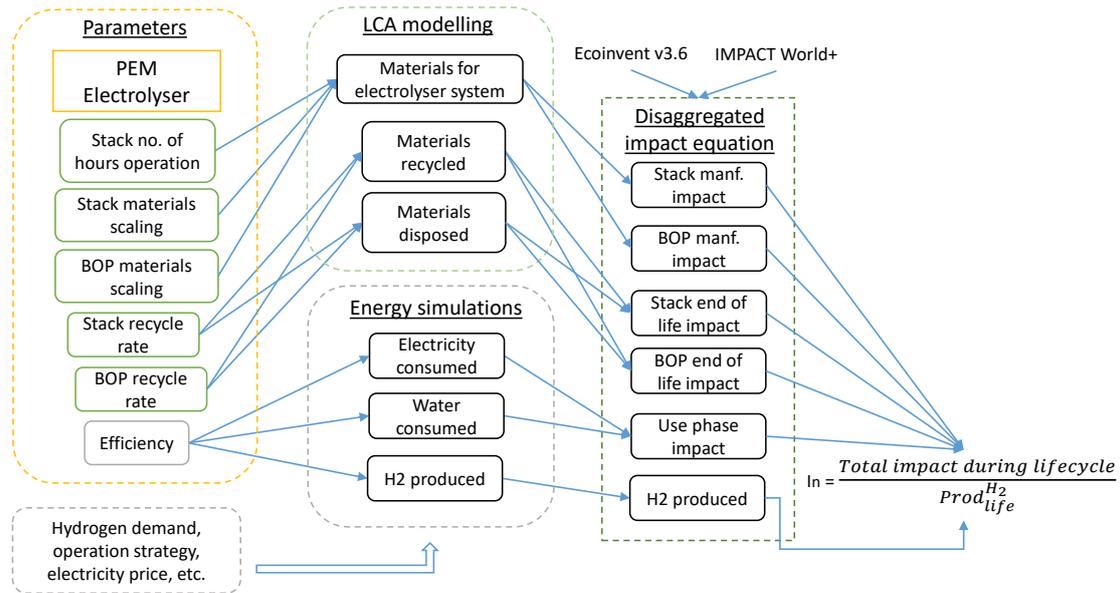


Figure 3.6: Description of the parameterized model for the PEM water electrolysis case study. New parameters added, that are relevant to LCA are illustrated in green.

3.3.5 Parameter definition for GSA

As mentioned earlier, the 6 selected parameters, presented in table 3.1 such that there is at least one parameter linked to the disaggregated equation inputs. Or when there is genuine uncertainty regarding them. To reiterate, the goal here is to find variability of the results as a function of parameters. Moreover as we saw in the previous subsection, since the parameters are directly related to each input of the disaggregated equation, the importance of the inventory data can also be estimated. since the first model is based on rapidly collected data, parameters or foreground processes can be preferentially searched. They are described below:

- The number of hours of stack functioning are limited because of degrading performance. Current suppliers report a range of 60k – 100k hours [Buttler2018]. The stack replacement frequency is decided based on an arbitrary decrease in performance (e.g. 10% decrease in efficiency relative to the start). Consequently this affects the amount of stack materials used.
- Efficiency of a PEM electrolyser is higher at when it is operating at less than its rated capacity [Kopp2017]. Even though we can estimate the operation using the energy simulation, it is based on electricity spot prices of 2014, hence we don't know the actual operating efficiency, neither its impact on the indicators. Additionally, there is degradation of efficiency as discussed above.
- Experts predict a decrease of up to 50% in the materials required for electrolysis in the future thanks to ongoing research efforts to increase cell area [Bertuccioli2014]. Therefore, relative variance of due to 50% decrease in the stack and BOP materials is selected. Also the robustness of the scaling for BOP can be verified.
- As mentioned earlier the recycling rate was approximated as 50% for the first iteration since it corresponds to the percentage of recycled materials used in metal manufacture [ADEME2018]. Also, no other parameter directly influences the end-of-life phase. Hence, recycling rate affecting the stack and BOP end of life is varied independently upto 95%.

Table 3.1: Input ranges to the DGSM sensitivity analysis

Parameter	Abbreviation used	Value	Source
Lifetime of stack	Lifetime	60,000 – 100,000 hours	[Buttler2018]
Efficiency of system (incl. auxiliary components)	Efficiency	50 – 60 %	[Buttler2018], [Kopp2017]
Multiplication factor for stack materials	Stack factor	0.5 – 1	[Bertuccioli2014]
Multiplication factor for auxiliary component materials	BOP factor	0.5 – 1	[Bertuccioli2014]
Recycling of stack material	stack recycle rate	50% - 95%	[ADEME2018]
Recycling of auxiliary component materials	BOP recycle rate	50% - 95%	[ADEME2018]

3.4 Case study: Results

This section proceeds as follows: In the first subsection we select the most influential parameters and corresponding foreground processes, supplied by the sensitivity analysis. Whereas in the second subsection, the influential parameters are precised or foreground process modelling is improved. These results are then compared to a conventional approach where all parameters or foreground modelling of a study were improved.

3.4.1 Sensitivity & Importance indices

The sensitivity indices (γ_e) for each one of the inputs in table 3.1 across the selected indicators are shown in figure 3.7. Higher the value of each parameter, higher is its influence on the final results. The two variables, efficiency and lifetime are selected since they are clearly the most influential variables. Amongst the rest, only stack factor is visible to have a share, but it stays less than 0.2 in all the indicators.

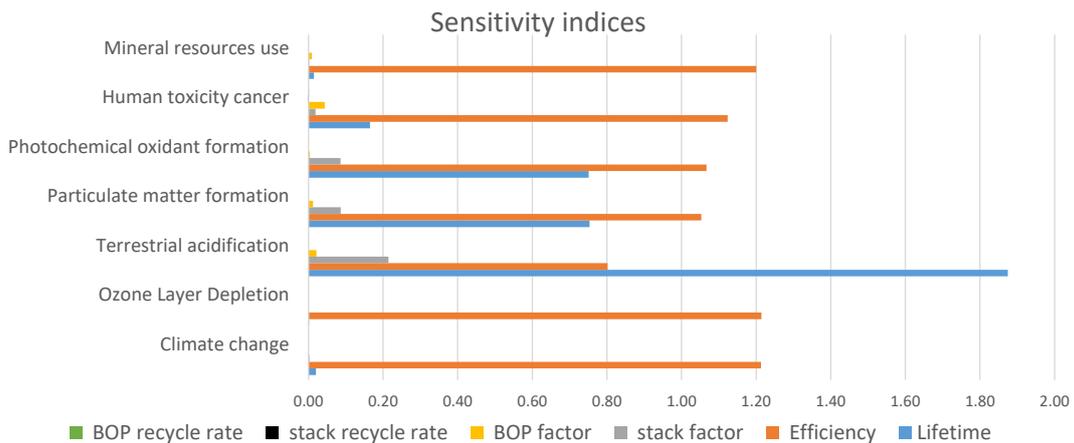


Figure 3.7: Sensitivity indices (γ) obtained for the parameters for each of the selected midpoint indicators

Furthermore, to calculate the Importance indices (Im), we can multiply the above sensitivity index with the corresponding contribution of the foreground processes. For instance, sensitivity index of efficiency for climate change, $\gamma_e^{climate\ change} = 1.21$. Efficiency is related to the use phase,

Table 3.2: Importance indices (Im) for 3 arbitrarily selected foreground processes from the use phase and the stack inventory. Full results in annex table A.6.

Foreground process	Classification	Importance indices (Im)						
		Climate change	Ozone Layer Depletion	Terrestrial acidification	Particulate matter formation	Photochemical oxidant formation	Human toxicity cancer	Mineral resource use
electricity, high voltage, production mix	Use phase	1.21	1.21	0.8	1.05	1.07	1.12	1.2
market for platinum	Manufacturing: PEM Stack	0.02	0.00	2.04	0.79	0.80	0.16	0.01
market for titanium	Manufacturing: PEM Stack	0.00	0.00	0.04	0.04	0.04	0.01	0.00

where the foreground processes are water and electricity consumption. Their respective contributions to climate change are $C_{use}^{elec.} = 99.94\%$ and $C_{use}^{water} = 0.06\%$. Hence, $Im_{use}^{elec.} = 1.21$ and $Im_{use}^{water} = 0.00$ for climate change. Values of these coefficients for 3 selected processes of each indicator is shown in table 3.2. A complete table with importance indices of the processes in use phase and stack inventory are included in table A.6.

It should be noted that, these indices are not a disaggregated representation of sensitivity indices. For example, the sensitivity index of efficiency for climate change, is not only affected by carbon content of the electricity consumed but also by the produced hydrogen (due to variation in the efficiency), see figure 3.6. But they can guide users to important foreground processes. Looking at the results from table 3.2, we select only the electricity and platinum processes, to be improved for the defined scope. Importance indices for the rest of the processes stayed below 0.02, similar to titanium.

3.4.2 Improving the influential aspects for the scope

In this subsection, the improvements of the identified influential aspects are described. In regards to the use phase, the parameter: efficiency and foreground process: electricity consumed are identified influential. For the efficiency, negligible improvements from the state-of-the art are anticipated at least until 2030. The principal reason is the comparatively higher capital costs than operational costs. As a result, research is focused on increasing current densities in the cells, which is detrimental to the efficiency [Schmidt2017]. Hence, an average efficiency of 57% is decided, that incorporates 3% degradation at the end of life from an initial efficiency of 60%. Whereas, the electricity mix is also identified in table 3.2 as a major contributing process to almost all impact indicators. Hence, the background modelling of French mix is updated to the expected mix in the year 2030 [ADEME2014]. The composition of the two electricity mixes is elaborated in the table A.7.

Lifetime is the other influential parameter, along with related foreground process of platinum. Experts predict stack lifetime of up to 85,000 hours [Schmidt2017]. However, some manufacturers already report a lifetime of 100,000 hours [Buttler2018]. In any case, within the above range, the number of stacks remains 2 for the case study, hence we fix the stack duration as 90,000 hours.

On the other hand, as identified by importance indices, platinum modelling was the other hotspot. This is improved as follows. Platinum was considered as a substitute for iridium in the first iteration, since the latter does not exist inecoinvent, but is a by-product of the same platinum group element extraction process [Ravindra2004]. Hence, iridium was modelled using the methodology of [Nuss2014]. In brief, the process ‘platinum group metal, extraction and refinery operations’ of ecoinvent v3.6 includes co-production of 5 elements: nickel, platinum, copper, palladium, rhodium and gold. The impact of each one is calculated according to economic allo-

cation. In the first step, this process was un-allocated using annual production values and their impacts, obtained from supporting files of ecoinvent database. Then, the impacts of 3 co-products of the same process, not included, (Rhodium, Iridium and Osmium) are reallocated according to the factors calculated by [Nuss2014]. These factors are obtained from economic allocation using historical market and production data. The allocation values of ecoinvent v3.6 and the ones used here are presented in table A.8. Furthermore, the amount of platinum and rhodium were updated according to expected reductions in 2030.

3.4.3 Results for scope 2030 and comparison with ideal case & literature values

In this section we compare the results for the three cases: base case, first iteration with easily available data; second case, where the important parameters and LCA processes are improved; ideal case, where even the non-influential aspects are improved. Recycling rate is increased to 95% as per [Zhao2018], where it is said high recycling is promised in the project contracts owing to financial values of rare metals. Material reductions in the stack inventory are updated for all elements according to the near future scenario of [Bareiß2019]. The differences between the 3 cases are outlined in table 3.3. The difference in the stack inventory can be seen in table A.1. The results for the LCA indicators are presented in figure 3.8, while keeping the base case i.e. the first iteration as the reference of 100%. The absolute values of three cases can be found in table A.10.

Table 3.3: Comparison of input parameters for the three cases

Parameter / Process	Base case	Important parameters adjusted	Ideal case	Source
Electricity mix	French mix 2016	French mix 2030	French mix 2030	[Wernet2016], [ADEME2014]
Recycle rate PEM stack & BOP	50%	50%	95%	[Zhao2018]
Efficiency	55%	57%	57%	[Buttler2018], [Schmidt2017]
Stack inventory	Current state of the art	Only platinum & iridium values adjusted. Iridium modelling improved.	'Near future' scenario	[Bareiß2019]
Lifetime stack	80,000 hours	90,000 hours	90,000 hours	[Buttler2018], [Schmidt2017]

From the figure 3.8 it is clear that adjusting the important parameters altered the results significantly. While further refining the model by changing all possible aspects of modelling has a negligible affect on the results. As the absolute error between the base case (first iteration) and ideal case ranged from 13% - 49%. Whereas when only the influential parameters are improved, the absolute error with the ideal case remains less than 1% across all impact indicators.

Thus in a real world scenario, the users or developers of the LCA models, do not need to search for data that were found to have negligible influence on the results. For instance, specifying the exact recycle rate for the year 2030 will only negligibly change the results. Thus saving time for the users. The same goes for the stack data, to obtain accurate estimation of impact, all of the PEM stack data do not need to be improved for the year 2030. Here the users or LCA practitioners could utilise their limited time to only accurately find the values of the catalysts platinum and iridium from literature, project collaborators or from other laboratories. Only the prediction of

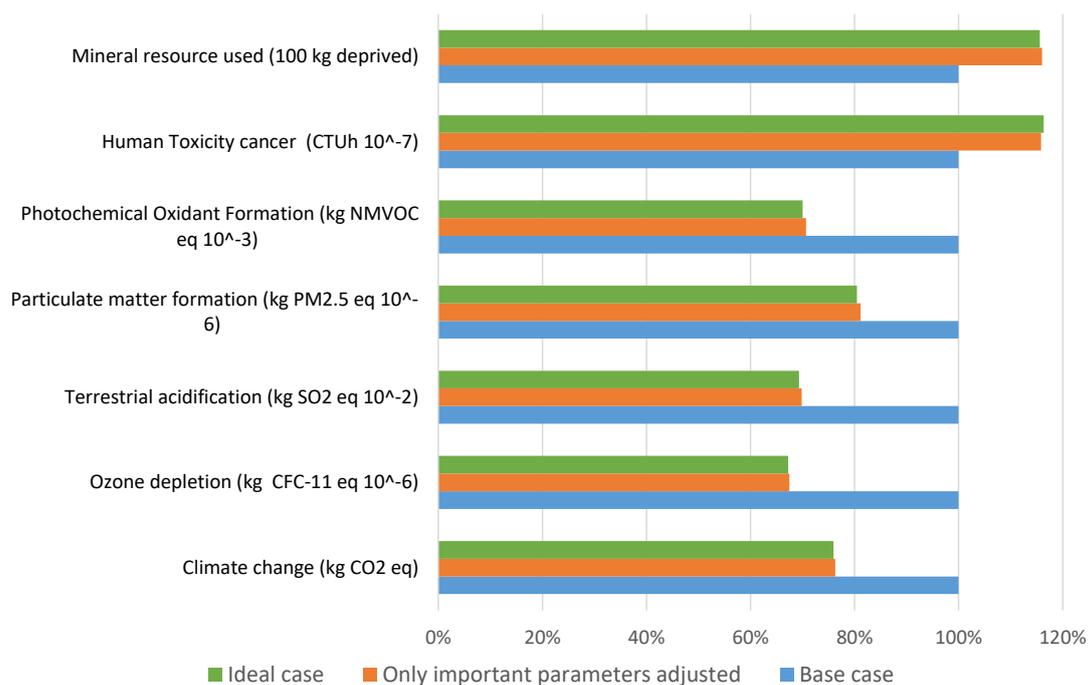


Figure 3.8: Results obtained for the LCA indicators for the three cases.

these values is sufficient to improve the model for the future.

One more additional insight using this model was regarding the material estimation of BOP materials. Even though they are auxiliary components, they are essential for the functioning of the system and require considerable amount of materials e.g. 13.7 tons of steel and 5.9 tons of copper (table A.2). However, even a 50% reduction in these materials were identified as negligible to the final impacts. Thus the scaling of these materials or future reduction in materials due to research, do not need to be improved for prediction of impact in the future.

Model consistency was verified using other unrelated investigations from the literature. The estimated French electricity mix of 2030 has a climate change potential of 55 g CO₂-eq/kWh, which results in a hydrogen production impact of 3.6 kg CO₂-eq/kg H₂. This is close to 3 kg CO₂-eq/kg H₂ estimated by [Bareiß2019], with electricity mix of ≈ 50 g CO₂-eq/kWh. Further, when the electricity mix is changed to only onshore wind, impact of hydrogen comes to 0.9 kg CO₂-eq/kg H₂. It lies in the range (0.5 - 2.3 kg CO₂-eq/kg H₂) of harmonised climate change impacts of hydrogen production from electrolysis using wind electricity calculated by [Valente2017]. Hence, the results obtained are coherent and agree well with the literature.

3.4.4 When influential parameters cannot be fixed

Often, it may happen that the variables related to highly influential parameters or processes cannot be fixed to a certain value. They could then be presented in terms of statistical values. In this section, an example is described by propagating and visualizing the uncertainty ranges. The distribution can be further analysed using statistical values. Let's choose a hypothetical case to evaluate this scenario. Say for our model, real-world efficiency and the electricity mix cannot be precisely predicted *a priori*. Reasonable efficiency range of 55% - 60% is selected. For the electricity mix, the three scenarios of the year 2050 proposed by [ADEME2014] are selected. The scenarios differ between them with variable share of nuclear energy with respect to renewable sources: wind, PV, hydroelectric and Marine energy. The actual mix in the future can lie between any of the 3 mixes. Hence a function is created varying the contribution of nuclear energy from 48%-18% with a proportional compensation by renewable sources, as per [ADEME2014]. Uncertainty is propagated

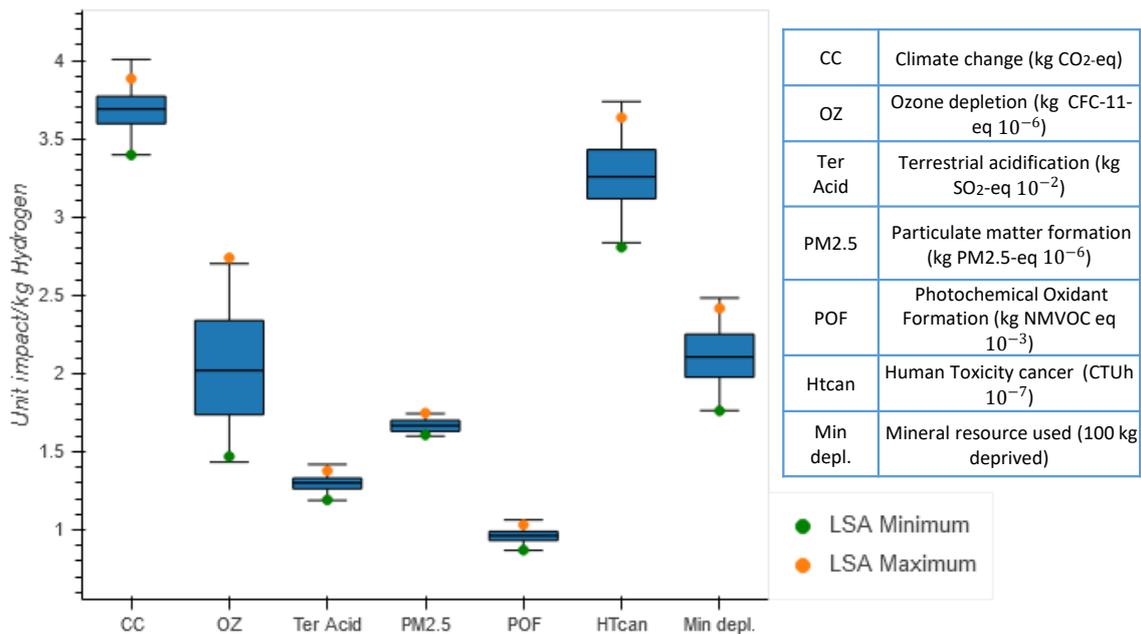


Figure 3.9: Boxplots of distribution of impacts obtained by propagating the PEM electrolyser system efficiency and electricity mix using quasi-monte carlo. Maximum and minimum values obtained by propagating extreme points of efficiency and electricity mix using LSA

using quasi-monte carlo or sobol sequence, which is also the precursor to DGSM. Both uncertain parameters are assumed to have an uniform distribution.

The results of the obtained distribution (300 samples) are illustrated in figure 3.9 using boxplots. For comparison, local sensitivity analysis (LSA) results are also plotted. Here, the extreme values of efficiency and electricity mix is propagated through the model to obtain maximum and minimum values for each impact. In boxplots, apart from the maximum and minimum values, space between each horizontal line represents approximately 25% of the values. For example, it can be said that for the considered uncertainty ranges, there is $\approx 75\%$ chance the climate change potential for producing 1 kg hydrogen remains below 3.8 kg CO₂ eq. If we compare this to local sensitivity analysis, only a range of global warming potential 3.39-3.88 kg CO₂ eq is obtained for the same uncertainties. Which is clearly lower as compared to the maximum value of 4.01 kg CO₂ eq that is actually possible. Thus outlining the utility of global sensitivity analysis.

To conclude, here even though a modest uncertainty range was used, there was still a difference between GSA and LSA. Thus in real-world situations, the possible minimum or maximum values of indicators can be significantly over or underestimated by local sensitivity analysis.

Furthermore, the preliminary version of this parameterized model was already verified to be working well with evolutionary algorithm for dimensioning [Sharma2020].

3.5 Conclusions

A step-by-step methodology is proposed to effectively establish parameterised LCA models for hybrid energy systems. The resulting LCA models are primarily developed to be used with energy simulations for the future context. The resulting LCA models thus take into account the alterations in the energy simulations and accordingly calculate the indicators. Thus a link between changes in design (system parameters, configuration) and LCA indicators was established which

did not exist before. The methodology was applied to case study involving hydrogen production from PEM electrolyzers. The parameterised model developed was connected to an energy simulation. The obtained results agreed well with unrelated investigations in the literature.

Prioritising data collection is proposed using global sensitivity analysis. Thus, when conducting an LCA with limited resources, the data which is the most important to improve the results could be identified using the proposed method. Taking into account the complexity with multiple uncertain parameters of an integrated model and the expensive computational time of hybrid energy simulations, use of derivative based global sensitivity measures (DGSM) is proposed as a GSA method. Along with a new co-efficient (Importance index) to identify the most important foreground processes. This particular type of GSA although used in other fields [Touzani2014, Kiparissides2009] has not been applied to life cycle assessment yet. The main advantage is that it offers good estimations of sensitivity indices most of the time (except for highly non-linear cases) while using only a fraction of the function evaluations as the state-of-the-art Sobol indices. Even in the worst case, it still allows reliable discard of non-influential parameters. This is really useful when the simulations are computationally expensive, there are many uncertainties in the model and exact sensitivity indices are not necessary, like in this context.

For this case study using the sensitivity and importance indices following data were found influential: system efficiency, electricity mix, catalysts amounts in the electrolyser stack and stack lifetime. Whereas recycle rate for stack and BOP, materials for BOP, stack materials other than catalysts were found not influential to the final impact. For the future context, even if the non-influential data was improved, it only resulted in less than 1% difference than the ideal case. Consequently, even in a highly uncertain context, the users do not have to accurately estimate the non-influential parameters or foreground processes correctly for robust estimations. For example in this case study, reductions in BOP or stack materials (except catalyst contents) for the future or detailed scaling of BOP materials is not necessary. Thus the framework can save significant amount of time.

Furthermore, quantified recommendations regarding the influential aspects could be provided to decision-makers. Using this information, for example, investment in R&D could be guided for improving the system efficiency or specifically reducing the catalyst content in the stack. Decreasing catalyst loading in the stack is also beneficial from an economic perspective [Schmidt2017]. On the other hand, research is not prioritised in improvement of stack efficiency due to economic reasons (ibid.). System efficiency is affected by stack efficiency and other auxiliary components. Whether this topic should be pursued is a different issue and beyond the scope of this thesis. The conclusion being that obtaining this type of insights are made possible.

The target audience envisaged for the methodology are personnel developing and using these software/models. Such a methodology did not exist in the literature thus far. It provides ideal starting points for developing these models, which are not included in the LCA handbooks. Especially starting with the literature search and how much data to collect. Followed by context specific, systematic recommendations on how to fill data gaps. Guidance is also provided for addressing the variabilities for sizing, geography and future evolution. Development of such models will enable integrated technical-economic-LCA assessments of hybrid energy systems. It therefore allows users or engineers to provide appropriate recommendations regarding R&D or planning of such technologies from an environmental point of view as well.

4

Objective reduction of LCA indicators using a visual survey

The chapter is commenced with the relevance of this proposal with respect to the literature followed by the current opinion about two types of impact indicators in LCA. Then the indicators preferred in the CEA organisation are discussed with a justification of choices made in this thesis. In the next section, the method of identifying redundant LCA indicators primarily for optimisation is then presented using a hypothetical design example. The limitations of this approach are outlined in comparison with other approaches found in the literature. Then in the next section the recommendations on its use are outlined. Finally, the chapter is concluded in the last section.

Contents

4.1	Introduction	80
4.1.1	Contribution to literature and relevance	80
4.1.2	Type of LCA indicators: midpoints and endpoints	80
4.1.3	Impact assessment method selected in this thesis	83
4.2	Indicator selection based on a visual survey	83
4.2.1	Example: Design of a bike	84
4.2.2	Design of a bike: comparison with PCA	87
4.2.3	Limitations	89
4.3	Recommendations on its use	90
4.4	Conclusions	90

4.1 Introduction

4.1.1 Contribution to literature and relevance

The importance of an optimisation approach including the environmental indicators along with the existing technical and economic indicators has been discussed in the first chapter. Objective reduction techniques are often used to improve the performance of multi/many-objective optimisation algorithms [Saxena2012]. Three reduction techniques used so far in LCA for this purpose were also briefly described. Relevance of using correlation-based objective reduction techniques even with the new generation evolutionary algorithm NSGA-III was also elaborated in section 2.3.5.

All of these methods, need at least a partial initialisation of the problem to obtain a sample set. Meaning that the problem has to be run at least for sometime to obtain an initial set of solutions. Furthermore, there could be problems with the sample set since the correlations in them might not represent the correlations between indicators on the pareto front [Yuan2017]. To avoid these problems, many investigations in the literature prefer to run the problem with all indicators for a long time. For example, around 16 hours in [Perez Gallardo2013] or 200-2000 generations in [Yuan2017]. This is actually counterproductive for objective reduction. Since, objective reduction is itself required to reduce the computational time.

In this thesis, a new technique to reduce the number of LCA indicators for dimensioning using a visual survey was found during the implementation of NSGA-III. It relies on identification of correlations between the LCA indicators using contribution analysis and is applicable to only a certain type of problems. Unlike the techniques presented, it does not require an initialisation of the optimisation problem. It only requires an estimate of impacts based on average simulation values. Since such preliminary calculations might be done by the engineers in any case, this alternative approach could save time. This approach is described in section 4.2 using an example of design of a bike. But first a background is provided on the LCA indicators in the following subsections.

4.1.2 Type of LCA indicators: midpoints and endpoints

Elementary flows obtained at the end of inventory analysis, are multiplied by characterisation factors in order to obtain quantified environmental impacts as outlined in equation 2.3. These factors are found by modelling impact pathways from the elementary flows to the areas of protection. Endpoints are located towards the end, quantifying the damage to the areas of protection [European Commission2010]. Three commonly defined areas of protection are: Human health, Natural environment and Natural resources.

On the other hand, the midpoints are located somewhere along the impact pathway. Ideally, they should be placed at the point after which the impact mechanism is the same for all flows [Hauschild2015]. For example, a simplified impact pathway of climate change is presented in figure 4.1. With the release of greenhouse gases, their concentration increases in the atmosphere increases. This leads to variable radiative forcing corresponding to a particular gas. Once this forcing is quantified, it is possible to commonly model the impact to the areas of protection. Thus it is a suitable midpoint indicator to quantify climate change. Impact pathways of impact indicators according to the impact assessment method ReCiPe 2016 can be seen in figure 4.2.

There has been a lot of discussion about when to use endpoints and midpoints in the literature but the debate remains open since both have their pros and cons [Bare2006, Hauschild2015, Pizzol2017, Bare2000].

Arguments for endpoints

Endpoints, owing to their modelling at the end of impact pathway, there are a lot more assumptions and value choices involved, leading to more uncertainties and decrease in robustness. Fur-

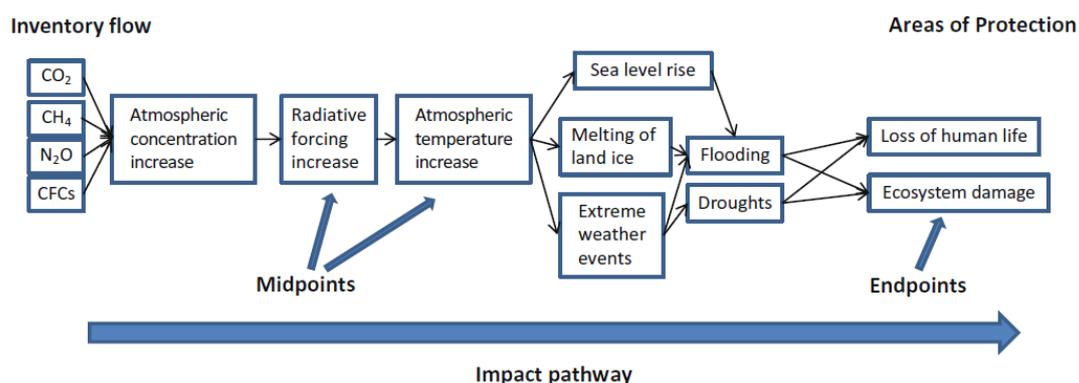


Figure 4.1: Simplified representation of climate change impact pathway from elementary flows to the areas of protection [Hauschild2015].

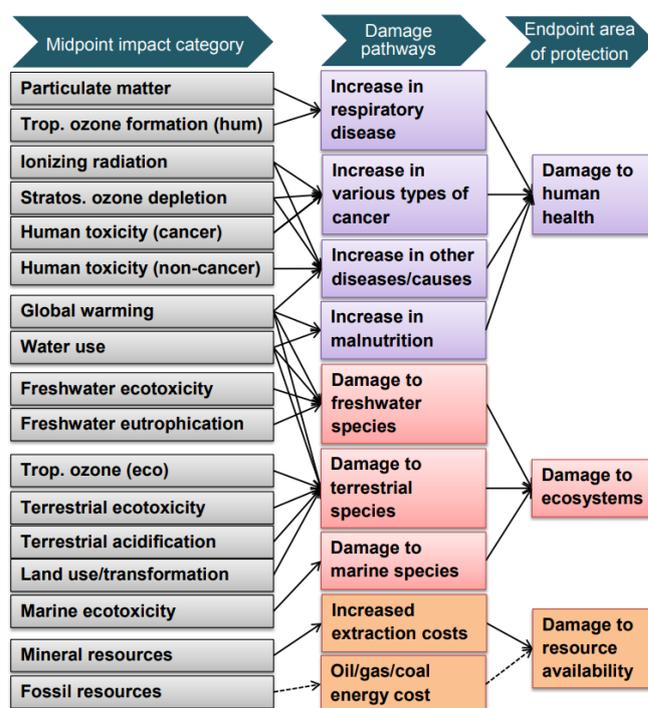


Figure 4.2: Impact indicators at endpoint and midpoints along with their impact pathways in impact assessment method ReCiPe 2016 [Huijbregts2016].

ther, midpoints being more comprehensive in the impacts calculated, not all of them could be included while modelling the damage pathways due to reasons such as lack of data or insufficient understanding [Bare2000].

There is a strong demand for policymaking based on LCA investigations, which in turn require simple and clear outcomes [Hellweg2014]. They conclude that this demand can be answered using endpoints or single score aggregation. In the 25th SETAC conference, a plenary session was organised to discuss the relevance of midpoints, endpoints and single point indicators for decision-making [Kägi2016]. It was discussed that even though endpoints have additional uncertainties and single score aggregation uses subjective weighting from experts, there is an argument that these two approaches might be, in fact necessary. The primary reason being, not all of the midpoints orient themselves in the same way. Consequently, to arrive to a conclusion about ‘whether product A is better than B’ or ‘Optimal way to improve a product/system’, decision-makers are forced to

subjectively attach weights to the midpoints or choose indicators based on their interests. In such a case, basing decisions on the endpoints instead might thus be a better approach.

Arguments for midpoints

On the other hand, midpoints are advantageous due to their comparatively low uncertainties. In terms of investigations quantitatively comparing midpoints and endpoints, the following studies were found.

Dong *et al.* compared the LCA results for construction materials according to midpoints and endpoints [Dong2014]. They show that the ranking of materials under both are generally consistent but when normalization is applied at each level, there can be considerable difference in the interpretation of results. That is, the relative magnitude of impact categories can be different depending whether the normalization is done at endpoint or midpoint. They recommend using midpoints and selecting the critical impact categories in the goal & scope. Whereas when endpoints are needed, midpoint results should also be provided.

Yi *et al.* compared outcomes from LCA investigation of municipal solid waste in Korea for midpoints and endpoints [Yi2011]. They concluded that if endpoints include all of the midpoints for its corresponding elementary flows and assumptions, the results are consistent. They further say that endpoints are helpful in interpreting midpoint results but a segregation of these impacts into global, regional and local will benefit this further.

Weidema compared the results from three endpoints methods Ecoindicator99, Stepwise 2006 and ReCiPe 2008 by monetizing them and subsequently revealing the underlying differences in their hypothesis and value choices [Weidema2015]. Lenzen *et al.* studied the consequences of using midpoints and endpoints for decision-making [Lenzen2006]. They conclude that even though endpoints offer more clarity in terms of the impacts, if the uncertainties are too high to answer the question, subjective judgements on more accurate midpoints or MCDA techniques could be used. Furthermore, they find that uncertainty in monetizing the damage can be as high as 80%.

Hauschild *et al.* identified the best practices amongst the midpoints and endpoints in the plethora of LCIA methods [Hauschild2013]. They concluded that most of the endpoints are in the need of further development before basing recommendations solely on them. It remains to be seen whether the updated methods since then (ReCiPe 2016 [Huijbregts2016] and IMPACT World+ [Bulle2019]) have improved these elements or not.

In general, midpoint indicators are largely preferred by the LCA practitioners since endpoint indicators can exhibit large uncertainties. Thus more robust conclusions could be derived using midpoints. Even when endpoints indicators or other types of aggregation are used, it is highly recommended to supplement them with the midpoint indicators [European Commission2010]. The disadvantages of using these aggregated methods were also discussed in section 2.3.3 where investigations have shown that they distort the pareto front of the original problem i.e. many problems that were optimal in the original set of midpoints could be skipped when aggregated indicators are used.

Preferred impact assessment in CEA

Midpoints are also preferred in CEA using the life cycle impact assessment method of the environmental footprint (EF) 3.0 [Fazio2018]. As compared to the ILCD recommendations in 2011 [European Commission2011], only midpoints are proposed under EF, where endpoints were also included initially. This method was proposed in order to facilitate communication of environmental information. Thus it is a part of the strategy of the European Commission to establish common methods to quantify and transmit environmental performance of different products. Consequently, the impact assessment methods proposed under the ILCD framework are further modified in this context.

4.1.3 Impact assessment method selected in this thesis

The impact assessment method selected in this thesis is IMPACTWorld+ [Bulle2019] due to three main reasons. First, it was recently updated using the state-of-the-art developments in the impact pathways (ibid.). Second, the EF method is not readily available in the LCA framework Brightway2 used in this thesis. While the IMPACTWorld+ is available for download online¹. Third, IMPACTWorld+ contains indicators at both midpoints and endpoints. It was not considered worth investing resources to add EF method to Brightway2. Mainly since, the impact assessment method if available, can be easily substituted and it will not make a difference to the methods proposed in this thesis.

4.2 Indicator selection based on a visual survey

This technique was discovered during the analysis of case-study results from the implementation of many-objective optimisation including the LCA indicators. It was realised that for certain, simple cases correlations between the LCA indicators could be found even before running an optimisation. That is, before an initialization of the problem which is required by other reduction methods. Where the problem is allowed to run for certain time and the results are analysed to identify correlations between the indicators.

For this method, correlations could be identified based on average values. After the correlations are identified, the objective reduction techniques operate on the same general principle. The correlated indicators are grouped together, resulting in uncorrelated groups of indicators. Then one indicator from each group is selected, such that each group is appropriately represented. Optimising with respect to the selected indicators will then be equivalent to when optimising with respect to the original set of indicators. Thus maintaining the same pareto front with the reduced set of indicators as the original problem. In other words, optimising with the reduced set of indicators should give the same solutions as optimising with the original indicator set.

Even though it has limited application, this type of pre-selection is not possible using the existing objective reduction techniques in LCA.

An outline of the method can be seen in figure 4.3 and can be described in the following steps:

First step is defining components to be optimised in terms of the defined functional unit as per the LCA goal and scope (impact per unit) for the considered set of indicators. For example, if the goal is to optimise electricity sources in a network, a valid unit of comparison could be impact per kWh or impact per kW. Similar for the case of LCA, the selection of the functional unit is crucial and can lead to significant difference in results. Nevertheless, since the technique mentioned here also uses the same functional unit for differentiation, results will be consistent with the LCA indicators.

In second step, impact per unit is then standardised by dividing with the corresponding maximum impact per unit. Since the impact indicators have different units, standardising enables comparison across them. At this point, impacts of components could be plotted on a line graph. If a particular component has a variable impact per unit, all possible values should be plotted.

Third step is grouping indicators such that each group has the same ranking of components within them. For different solutions consisting of different combinations of components, indicators

¹ <http://www.impactworldplus.org/en/writeToFile.php>

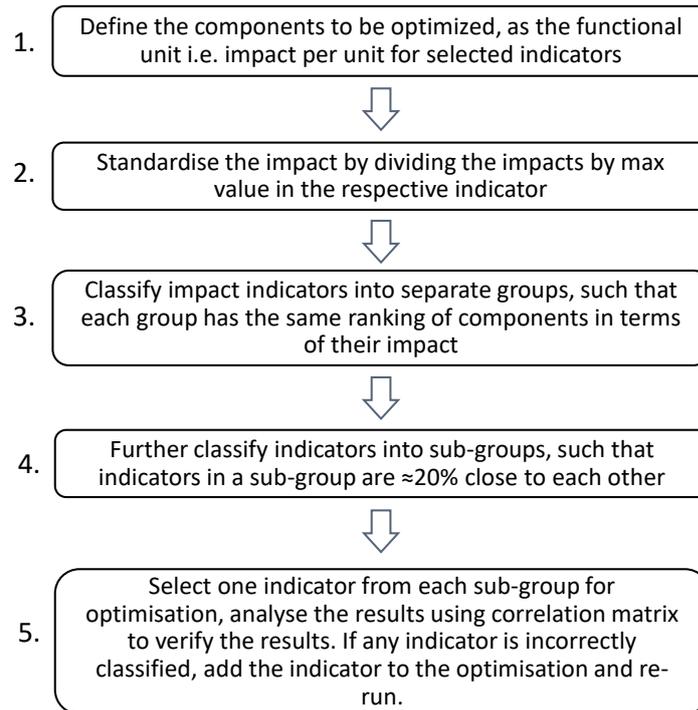


Figure 4.3: An illustration of the methodology to select impact indicators using a visual survey

within a particular group will co-vary. Since the impact of the system is nothing but an aggregation of impacts of individual components.

Fourth step is then further sub-grouping the indicators such that the indicators in a sub-group lie close to each other. It was seen in the simulation runs that a range of $\approx 20\%$ could be used to distinguish the sub-groups. This step further refines the groupings obtained in the previous step. It gives a more conservative classification of indicators since it was observed during the analysis that in certain cases even if the indicators belong to the same group (same ranking), they could vary differently if they lie far apart from each other. One indicator from each sub-group could then be selected for optimisation.

Fifth step is about verifying the obtained results. The indicators from the obtained optimisation results could be verified using a correlation matrix. The indicators within a particular sub-group should have high coefficient of correlation (≥ 0.7). If this is not the case, this indicator should be included in the optimisation and the problem should be re-initialised.

4.2.1 Example: Design of a bike

For a demonstration of the technique, let's consider a hypothetical problem of designing a bike weighing 1 kg. The functional unit selected for LCA is also 1 kg for simplicity. It is made from 3 materials: aluminium, glass fibre and tin. The environmental impacts of these 3 materials are obtained from ecoinvent 3.6 [Wernet2016] using the ImpactWorld+ assessment method [Bulle2019]. Complete name of the processes and absolute impact per kg is presented in annex table A.11.

Following the first step of the methodology, the functional unit 1 kg is the comparison unit between the 3 materials. Then, the impact of these 3 materials is standardised with respect to the maximum value and plotted on a line graph. For example, aluminium has the maximum value in climate change potential, thus its value becomes equal to 1. This enables the comparison across different environmental indicators. These 8 midpoint indicators out of the 18 are arbitrarily

selected in this case. Application of the methodology to this case study for 1-4 steps can be seen in figure 4.4.

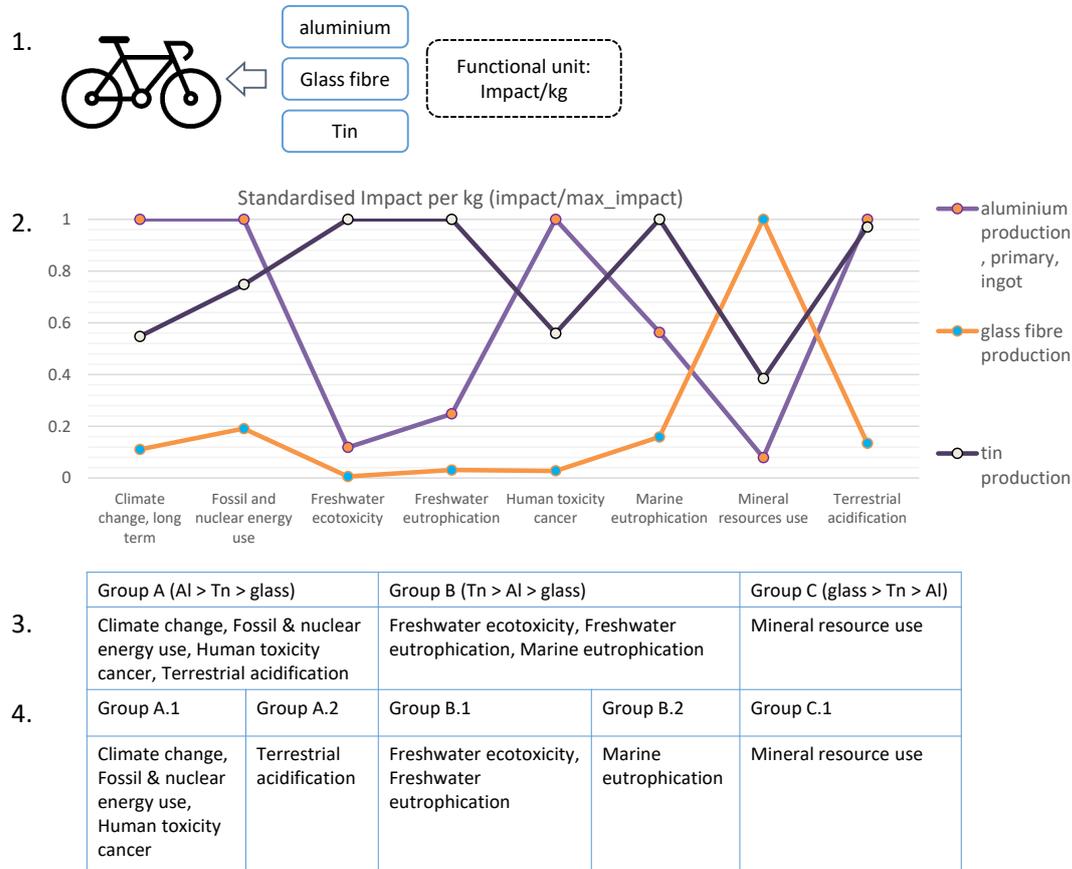


Figure 4.4: Classification of the indicators according to the visual survey (steps 1-4) applied to a fictional case study of designing a bike

Then for the third step, indicators can be classified into 3 groups, with respect to the ranking of the 3 materials within them. For example, in climate change, highest impact is from aluminium, followed by glass fibre and tin.

The indicators are further classified into sub-groups such that they lie close to each other (within $\approx 20\%$) in the fourth step. Thus 5 sub-groups are obtained as seen in the figure. Optimisation is not done in this case, but 20 design alternatives are generated with varying quantities (in kg) of the 3 materials. These alternatives are generated by increasing the aluminium values and then filling the rest of weight with glass fibre and tin. Their compositions and impacts, standardised with respect to the max value is shown in figure 4.5. It can be clearly seen that the indicators within a group and a sub-group co-vary together. While the co-relation between the sub-groups is even stronger.

To understand in more detail, let's consider the two indicators belonging to the same sub-group, freshwater ecotoxicity and freshwater eutrophication. The impact per kg in decreasing order in them is: aluminium, glass-fibre and tin (figure 4.4). The relative positioning of points within them is quite similar. A particular solution of designing a bike weighing 1 kg will require x, y and z amounts of aluminium, glass-fibre and tin respectively. For the two indicators, each material will have a similar impact contribution, since the relative impacts of the three materials in both indicators are similarly positioned. Consequently since the impact of the system is nothing but an aggregation of the impacts of 3 materials, the indicators within a subgroup will thus co-vary.

If we compare 5 solutions in terms of their contributions in figure 4.6, after having standardised

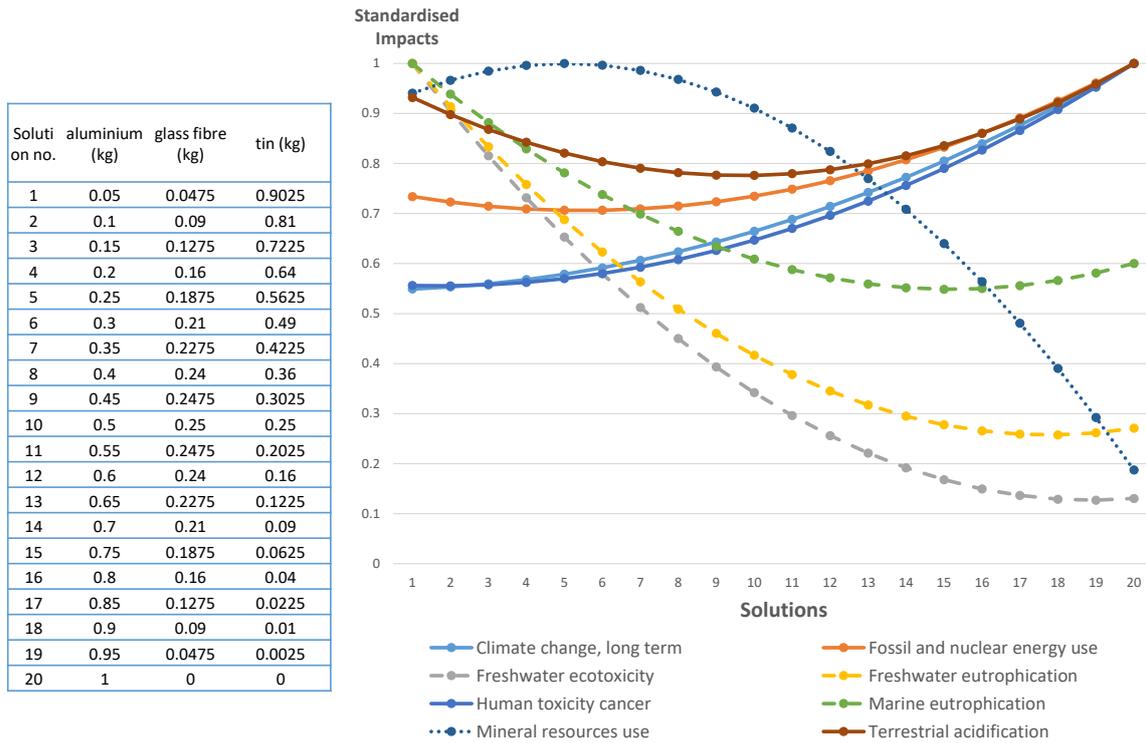


Figure 4.5: Results of impact indicators of the 20 design alternatives

them with respect to the max value in the set, the relative change and contribution analysis between the two indicators is similar. Since in a particular sub-group, ranking of these materials in terms of impact is similar, the impacts within a classified sub-group will also change proportionally. Even though the percent decrease is different in both figures, the co-efficient of correlation and co-efficient of determination for the two indicators are ≈ 0.99 . Thus, outlining a perfect correlation between both. The efficiency of classifying co-varying indicators can be verified using a correlation matrix in the next section.

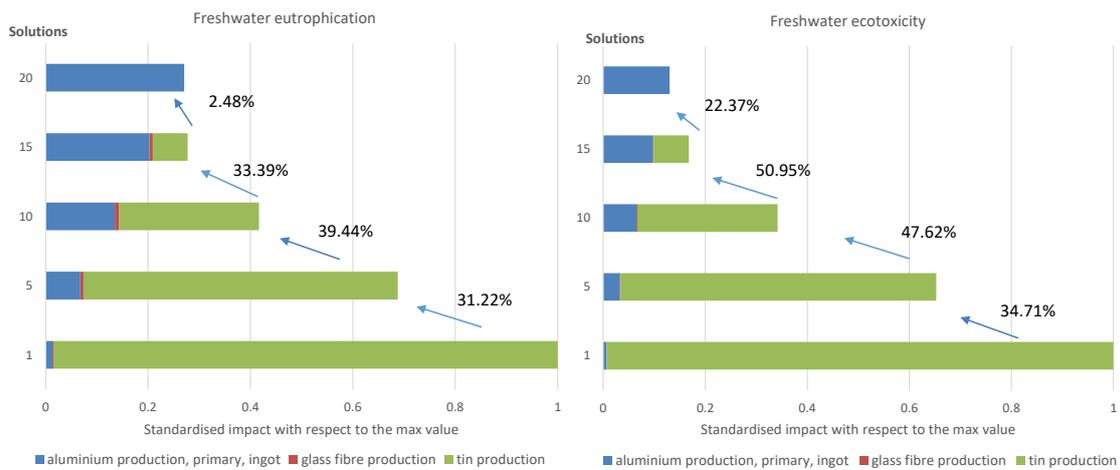


Figure 4.6: Contribution analysis of the 5 solutions in terms of the two LCA indicators.

4.2.2 Design of a bike: comparison with PCA

The selection of indicators using the visual survey can be compared to selection using principal component analysis. The technique was initially described in section 2.3.4. To remind the steps to be followed in this technique, they are illustrated in figure 2.10 as per [Deb2005, Saxena2012].

Here the sample set for PCA based reduction are the hypothetical 20 design alternatives presented previously. The analysis was done using a freely available excel package Real statistics² by [Zaiontz2017]. Once the solutions are available, the next step is the correlation matrix which can be seen in table 4.1.

Table 4.1: Correlation matrix of 8 impact indicators for the 20 bike alternatives.

	<i>Climate change</i>	<i>Fossil and nuclear energy use</i>	<i>Freshwater ecotoxicity</i>	<i>Freshwater eutrophication</i>	<i>Human toxicity cancer</i>	<i>Marine eutrophication</i>	<i>Mineral resources use</i>	<i>Terrestrial acidification</i>
Climate change	1.00	0.97	-0.86	-0.84	1.00	-0.74	-0.98	0.53
Fossil and nuclear energy use	0.97	1.00	-0.72	-0.69	0.98	-0.56	-1.00	0.71
Freshwater ecotoxicity	-0.86	-0.72	1.00	1.00	-0.84	0.98	0.74	-0.02
Freshwater eutrophication	-0.84	-0.69	1.00	1.00	-0.81	0.99	0.71	0.02
Human toxicity cancer	1.00	0.98	-0.84	-0.81	1.00	-0.71	-0.99	0.57
Marine eutrophication	-0.74	-0.56	0.98	0.99	-0.71	1.00	0.59	0.18
Mineral resources use	-0.98	-1.00	0.74	0.71	-0.99	0.59	1.00	-0.69
Terrestrial acidification	0.53	0.71	-0.02	0.02	0.57	0.18	-0.69	1.00

Then, the eigenvalues and eigenvectors of this matrix is calculated. The resulting eigenvector corresponding to the highest eigenvalue is the first principal component (PC). These eigenvectors are arranged in decreasing order in table 4.2. Furthermore, 100% of the cumulative variance is explained by the first two components. These results can be seen in figure 4.7. Following the recommendations of [Saxena2012], most negative and most positive indicators are retained. These 3 (out of 8) indicators are highlighted in bold in table 4.2. Meaning that just these 3 indicators could be used for optimisation while avoiding a transfer of impact. Or in the optimisation terminology, optimising with respect to these 3 indicators will maintain the pareto structure of the initial 8 indicators. Since the variance in the entire set can be in fact predicted by estimating only these 3 indicators. They cannot be further reduced since the coefficient of correlation between them stays less than 0.53 for all combinations.

² <https://www.real-statistics.com/>

Table 4.2: PCA results for the 8 impact indicators. Selected indicators are highlighted in bold.

	PC ₁	PC ₂	PC ₃	PC ₄	PC ₅	PC ₆	PC ₇	PC ₈
Eigenvalues (λ)	6.23	1.77	0.00	0.00	0.00	0.00	0.00	0.00
Explained variability (%)	78	22	0	0	0	0	0	0
Cumulative variability (%)	78	100	100	100	100	100	100	100
Climate change	0.399	0.077	0.265	-0.170	0.015	-0.331	0.780	0.133
Fossil and nuclear energy use	0.377	0.253	0.362	-0.331	-0.181	-0.425	-0.566	-0.139
Freshwater ecotoxicity	-0.365	0.311	-0.005	-0.204	-0.427	0.067	0.267	-0.686
Freshwater eutrophication	-0.357	0.339	0.016	-0.233	-0.457	0.044	0.000	0.702
Human toxicity cancer	0.396	0.116	0.288	-0.206	-0.028	0.839	0.000	0.000
Marine eutrophication	-0.324	0.442	0.099	-0.338	0.759	0.000	0.000	0.000
Mineral resources use	-0.382	-0.230	0.840	0.310	0.000	0.000	0.000	0.000
Terrestrial acidification	0.175	0.676	0.000	0.716	0.000	0.000	0.000	0.000

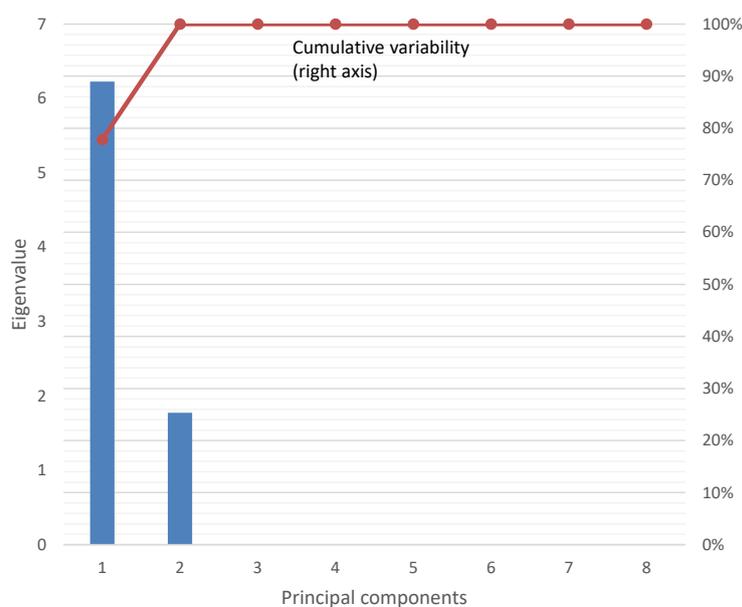


Figure 4.7: Scree plot of the principal components

The impact indicators are shown in bi-dimensional plots of the first two principal components in figure 4.8. This plot makes it easier to visualise the relationships between the different indicators. The closer the indicators lie to each other, the more they are co-related and vice-versa. The co-variance predicted by the visual survey is thus corroborated, first by the correlation matrix and then by the figure in discussion.

In terms of selecting the impact indicators for optimisation, it is clear that PCA is more efficient than the visual survey which can be seen in terms of two things. First, the number of indicators selected by PCA is 3 whereas by visual survey is 5. The latter is more conservative in this aspect. Secondly, PCA can identify indicators with the maximum and minimum loadings on the respective PC. That is, quantification of impact indicators with the highest influence within a group. Even though selection of an indicator within a sub-group is random for the visual survey, its effect on the final optimisation result should be negligible since it was shown that reliably grouping of

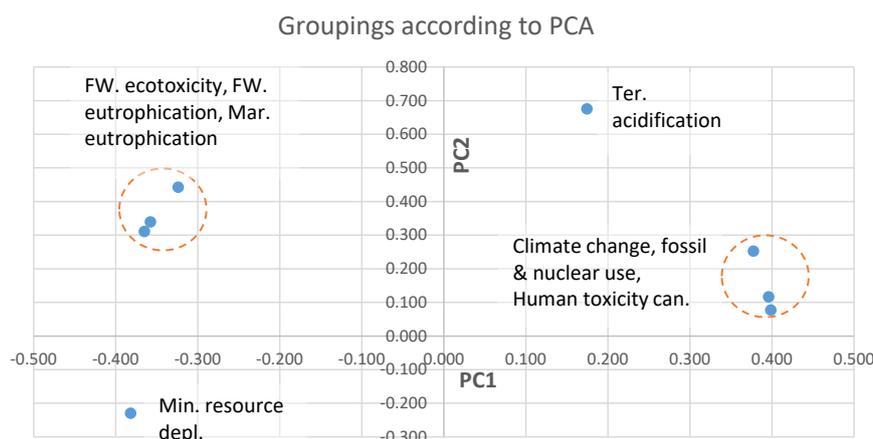


Figure 4.8: LCA indicators plotted in terms of the two principal components.

Table 4.3: Comparison of indicators selected by visual survey and PCA

Initial indicators (8)	Climate change, Fossil & nuclear energy use, Freshwater ecotoxicity, Freshwater eutrophication, Human toxicity cancer, Marine eutrophication, Mineral resource depletion, Terrestrial acidification
PCA reduced indicators (3)	Climate change, Mineral resource depletion, Terrestrial acidification
Visual survey indicators (5)	Climate change, Mineral resource depletion, Terrestrial acidification, Marine eutrophication, Freshwater eutrophication

indicators in terms of co-variance was achieved.

4.2.3 Limitations

Following are the limitations of identifying redundant indicators using the visual survey approach:

When the optimisation variable cannot be defined in terms of functional unit - In the above example each optimisation variable (amount of material) could be defined with respect to each indicator (impact per kg). In case when this is not possible, a visual survey cannot be used. For energy systems, a parameter related to operation strategy is a commonly used optimisation variable. For instance in example 1.1.5, deciding the value of electricity price ' \hat{x} ', such that, the electricity is taken from the grid or from local sources. It is not possible to define the optimisation variable electricity price in terms of the functional unit (impact per kg H₂), thus making visual survey unusable.

Complex system - A sufficiently complex system could render the use of visual survey infeasible since it might not necessarily reduce the number of indicators to optimise. The number of sub-groups formed following the schematic in figure 4.3 could be equal to the original set of indicators. Two possible scenarios come to mind. First, if there are too many components, there could be more than manageable subgroups due to numerous rankings possible between the components. Second, if the components are highly non-linear in the line graph (e.g. figure 4.4).

Conservative reduction of redundant indicators - As compared to other methods, visual survey can be more conservative in selecting the representative indicators. With this approach it could still exist the opportunity to further identify redundant indicators. However, it was

seen in various simulation runs that if the classification in the visual survey was stopped only with the technology rankings (step 3 in figure 4.3), consistent identification of correlations might not be always guaranteed. Hence, it is obliged to sub-group them further and select additional indicators.

4.3 Recommendations on its use

Taking into consideration the above discussion, the visual survey technique is thus recommended to be used when: a case is not complex and the optimisation variables can be defined in terms of the LCA functional unit. The effort saved in this case will be the initialization of the problem for generation of sample set. Usually, it is running the problem in the original objective space for a limited time. All the other strategies seen so far need at least a partial initialisation of the optimisation problem to obtain a sample set. As mentioned previously, the problem is often run for many generations (e.g. 200-2000 generations in [Yuan2017, Saxena2012, Perez Gallardo2013]) to obtain this initial set so that it represents the same correlations/dominance structure as in the pareto front. For situations where there are many objectives or when the model is expensive to compute, this can take a lot of time. Thus making its application impractical. On the other hand, the visual technique seems to be the only approach which could be utilised '*a priori*'.

Additionally, the analysis of an optimisation problem in this way (e.g figure 4.5) offers an intuitive approach to finding correlations and objective reductions. Thus enhancing the understanding of the model behaviour. This can be especially relevant when the users are non-experts in life cycle assessment or objective reduction.

More research is needed to improve the above mentioned limitations. The most important step would be the application of this principle in an algorithm. Using this, the technique can be refined and the practical ceilings on its use can be found. For example, for forming the sub-groups of indicators in step 4 of the method, a 20% limit between the indicators is proposed using empirical results. This number could be tested or further defined precisely. Similar testing is needed for finding insights on its application on complex systems. For instance, further precision of use on components with non-linear impact/unit.

4.4 Conclusions

A visual technique to identify correlations between the LCA impact indicators for certain systems was presented. This property could then be leveraged to select impact indicators for optimisation and discard the redundant ones. A hypothetical case study of designing a bike was used as a demonstration. An already well established approach in the literature (PCA) for the same purpose was used to verify the obtained results. The limitations of this approach and its comparison with similar found techniques in the literature were also discussed. It is concluded that it is a promising technique that can identify correlations between the indicators without an initialization of the problem. It could be used in certain cases if the limitations can be avoided. Nevertheless, more research is required in order to implement the principle of this technique in an algorithm to refine its applications and also to eliminate human errors.

5

Including environmental indicators in the multi-objective search of solutions

The goal of this chapter is to present an optimisation approach of combining NSGA-III and objective reduction to include high number (10+) of LCA indicators in the multi-objective search of solutions of hybrid energy systems. This chapter proceeds as follows. First the contribution and relevance of this proposition with respect to the literature is discussed. Then in the second section a dimensioning case study involving hydrogen production from PEM water electrolysis is prepared for demonstration of this approach. The proposed optimisation approach is tested with 14 indicators out of which 13 are LCA midpoints in the third section. Its performance with respect to the conventional NSGA-III and multiple extents of objective reduction is presented. Then in the fourth section an approach to visualise pareto-optimal solutions at the end of optimisation is presented. Conclusions are presented in the last section.

Contents

5.1	Relevance and contribution to the literature	93
5.2	Hydrogen production case study	94
5.2.1	Inputs to the integrated model	95
5.2.2	Computational structure	97
5.3	Performance of NSGA-III with reduced indicators	98
5.3.1	Hypervolume indicator	99
5.3.2	Indicator reduction using visual survey	100
5.3.3	Indicator reduction using PCA	102
5.3.4	Reference point specification	104
5.3.5	Results with NSGA-III and reduced indicators	105
5.3.6	Discussion	110
5.4	Visualization of solutions for decision-making	112
5.4.1	Proposed procedure	112
5.4.2	Application to the obtained results	113
5.4.3	Comparison with the status-quo	117
5.5	Conclusions	118

5.1 Relevance and contribution to the literature

For the inclusion of LCA indicators, so that acceptable solutions can be searched with higher environmental efficiency, NSGA-III algorithm with objective reduction is proposed. Objective reduction techniques indeed improve performance of various evolutionary algorithms in terms of lower computational time and better quality results in the literature [Brockhoff2007, Gonzalez-Garay2018, Deb2005, Perez Gallardo2013]. To the best of our knowledge, objective reduction techniques are not yet combined with the NSGA-III. A possible reason could be that objective reduction techniques were initially introduced to deal with the convergence issues of the older generation evolutionary algorithms like NSGA-II or SPEA2. While NSGA-III already seems to be efficient with high number of objectives (upto 15) [Deb2013]. However, we saw in section 2.3.5, that promising improvements on computational time, solution quality and decision-making could be possible using objective reduction especially when the computations are expensive.

As far as it is understood, the evolutionary algorithm NSGA-III has not been utilised in LCA investigations thus far. Except for conference article that was made with the preliminary model of the case study in this chapter [Sharma2020]. With its demonstration, two key advantages are brought to the LCA literature. First is its ability to supply well converged, well-diversified solutions for problems with up to 15 objectives [Deb2013]. Second is its ability to search for preferential solutions using the supplied reference points. In the LCA literature, the pareto solutions are often subject to post-processing using techniques such as multi-criteria decision making [Perez-Gallardo2018], data envelopment analysis [Gonzalez-Garay2018], pareto filters [Antipova2014], etc.. Using NSGA-III, preferential solutions are inherently searched by supplying a set of reference points. Thus, use of such post-processing techniques might not be necessary.

In this chapter, a methodology to include LCA indicators in the many-objective optimisation of hybrid energy systems is proposed. The performance of NSGA-III is thus tested with and without objective reduction on a dimensioning problem in hybrid energy systems for the original 14 objectives. Visual survey and PCA, both correlation based techniques, were used for objective reduction. It was found that better quality solutions could be obtained in less time when the objectives were reduced until 7. If objectives were reduced further using a correlation matrix as per the last step (see figure 2.10) suggested by [Deb2005, Saxena2012], there was a trade-off between the computational time and the quality of solutions.

Furthermore, the correlation between indicators was also used to ease visualization of results. Solutions could be presented only in terms of reduced indicators along with the correlations information of the temporarily eliminated indicators. Such that, the decisions could be made on the reduced set while clear information is provided regarding the their correlations with the discarded ones. Here on the other hand only 4 indicators were found to be sufficient for the final representation from the initial set of 14 optimisation objectives.

This proposal thus addresses the problem of including the LCA indicators to enable the search of solutions with higher environmental efficiency and supplying them in a coherent way to engineers or decision-makers. Additionally, 13 midpoint LCA indicators are included for this purpose instead of aggregating them using endpoints or weighting. This addresses two issues: First, the solutions optimal in the original objective space are not skipped. Second, it also adheres with the preferences of LCA community in general, decision-makers and CEA & its partners of using midpoint indicators.

5.2 Hydrogen production case study

The case study of hydrogen production presented in the earlier chapters is modified in order to demonstrate the optimisation approach. Basically, the problematic, constraints and the goal remain unchanged but two additional components, photovoltaics and wind turbine are added. For a reminder, the goal is to dimension the system in order to satisfy the variable hydrogen demand in Marseille, France. Hydrogen is produced using PEM water electrolysis. The restraints to this sizing problem are costs (fixed and variable) of each component, energy demand and local weather conditions. An illustration of the case study in the simulation software can be seen in figure 5.1.

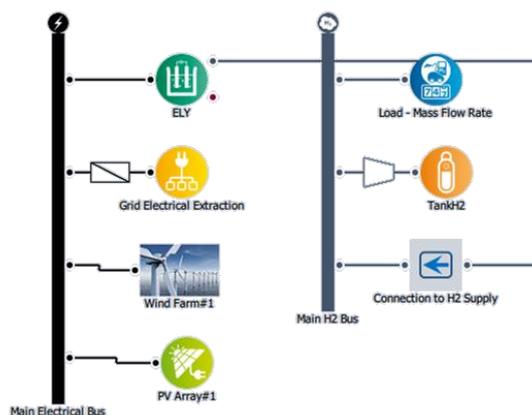


Figure 5.1: Representation of the case study in Odyssey.

Pre-estimation

Before proceeding with the optimisation, a pre-estimation is necessary to see whether it is actually relevant for this case. If the difference between the min and max values of LCA indicators is less than 30%, due to the uncertainties in LCA, it will be difficult to differentiate between the obtained solutions. Thus an optimisation search would not be very useful. For the above case, a pre-estimation was already made in an example in section 1.3.2, where the inclusion of LCA indicators in the multi-objective search was justified.

It was seen that only considering the impact of electricity, the climate change impact per kg H₂ reduced from 4.22 kg CO₂-eq (grid electricity in France) to 0.83 kg CO₂-eq (electricity from wind turbine) respectively. Thus there could be a substantial gain in the environmental efficiency. This is again a rough estimation and the final impact will depend on the system constraints, local weather conditions, end of life scenarios, etc. However, it gives enough motivation to proceed forward with this dimensioning problem. Thus in this chapter, the system is dimensioned with respect to the technical, economic and LCA indicators together.

Modelling information

The technical, economic and LCA inputs are described in detail as much as possible in the coming subsection. The tech-eco indicators are retained from the example presented in chapter 1. Techno-economic data was found from literature, databases or default values used in the laboratory. For developing the LCA model, the framework presented in chapter 3 is used again. For simplicity, these steps are not described in terms of the framework. Instead the information is described according to each component of the system.

The technical indicator is the unsatisfied hydrogen demand, which can be expressed as the percentage of the total mass based hydrogen demand the system was unable to satisfy during the project lifetime. While the economic indicator is levelised cost (€) per kg of hydrogen. While for LCA as mentioned before, 13 midpoint indicators from a recently updated method IMPACT-World+ are selected [Bulle2019]. The main reason for this selection was that they were recently updated according the developments in the impact pathways. Any other impact assessment could also be used in the future by the users.

5.2.1 Inputs to the integrated model

The economic inputs are only for demonstration of the approach and they are based on default values found in the literature. Project lifetime is assumed to be 20 years, as it is the often used value for similar energy projects. The step 3 of GSA is not performed for this case study. It was however already performed for an initial version in the earlier chapter. A summary of the model is presented in table 5.1 and it is further described below:

Electrolyser data is the same as the one established in the previous chapters for the 'base case' of PEM water electrolysis. Except for one difference which is, the modelling of iridium process not available in ecoinvent 3.6 was added according to the method from [Nuss2014] described previously. The stack life time is assumed to be 90,000 h. Full data is available in annex tables A.1 (base case), A.2.

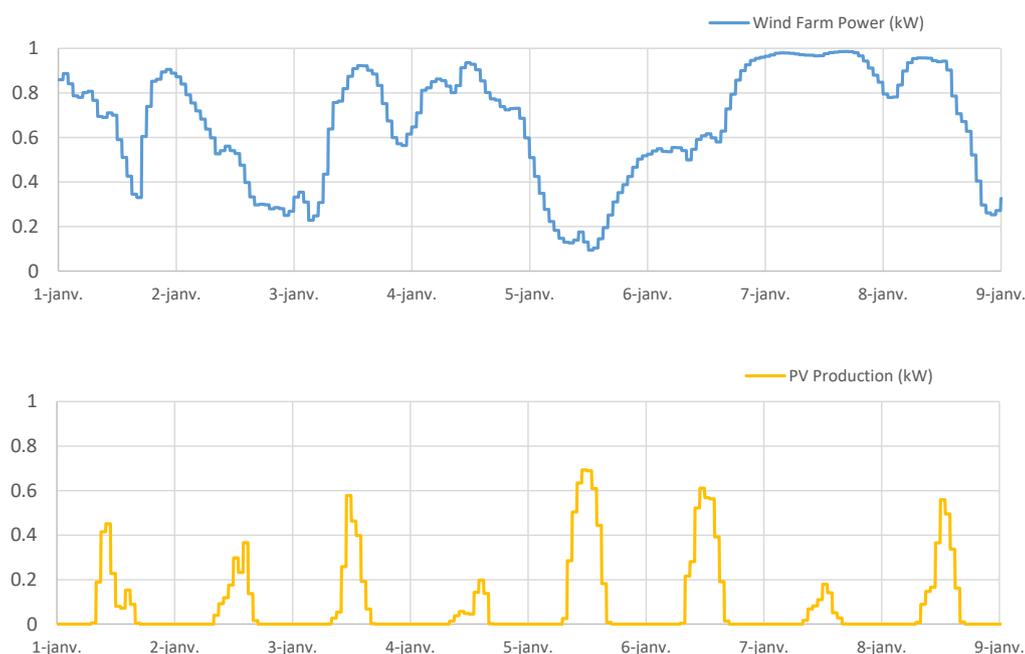


Figure 5.2: Production data of solar PV and Wind for the first nine days.

Solar PV production data is taken from the freely available database renewable.ninja¹ [Pfenninger2016]. For the location of Marseille, France, the capacity factor is 0.19. The LCA model for solar PV is taken from the ecoinvent 3.6 database for multi-Si technology, updated according

¹ <https://www.renewables.ninja/>

to more recent industrial data from [Gazbour2016]. More specifically, four elements are modified. First, the wafer thickness is decreased according to recent modules, resulting in a decrease in amount (kg) of silicon per 1m^2 wafer manufacture from 1.14 to 0.66. Second, in multi-Si silicon production process, only solar grade silicon is used. Third, in the silicon production, siemens process, electricity requirement is decreased from 115 to 55 kWh/kg. While the electricity mix is changed to 80% chinese and 20% world mix, inline with the current solar PV supply chain [Gazbour2018]. An illustration of the modifications in the ecoinvent 3.6 database is illustrated in figure A.1 and the PV inventory activities are listed in table A.9. Solar PV is a modular technology, where the size is increased with additional PV panels. Thus, a linear relationship between the solar PV size and the inventory is followed i.e. $b=1$ in equation 3.2.

Wind production data is taken from renewable.ninja as well [Staffell2016]. The data was extracted for a fixed hub height of 76m which corresponds to 2MW turbine with a rotor diameter of 91m. This results in a capacity factor of 0.31. For other wind turbine sizes, capacity factor is kept constant, that is, the time series is scaled linearly with respect to the turbine power. In the real world however, the wind speed and thus the electricity production varies with the hub height (which changes with turbine size). Another approach of using a shear co-efficient as done in [Caduff2012] was also tried, but it resulted in even more unrealistic estimations. The ideal case is extracting data for each turbine power and hub height from renewable.ninja itself. However, they only allow limited access per hour which could be a limiting factor for sensitivity analysis or optimisation. This was not pursued further. Hence, there is room for improvement in the electricity estimations from wind turbine. The LCA inventory is from [Padey2013, Sacchi2019], available online² for adaptation as a python notebook. It is a parameterized model that calculates a detailed inventory using a limited amount of parameters such as turbine size, type (offshore or onshore), etc. An example inventory of 2MW wind turbine calculated from this model is shown in annex B.

Table 5.1: Summary of inputs to the integrated techno-economic-LCA model

Component	Technical	Economic		Environmental
		Initial fixed costs	Other costs	
Electrolyser	56% efficiency stack life - 90,000 h	3.4 €/size (W)	Stack change: 2% of initial cost	Parameterized model, data in base case table A.1 & table A.2
Solar PV	5 min time-serie Cap. factor - 0.19	0.8 €/size (W)	O&M: 0.42% of initial cost	Ecoinvent 3.6, update acc. to [Gazbour2016]. figure A.1, table A.9
Wind	5 min time-serie Cap. factor - 0.31	1.38 €/size (W)	O&M: 3% of initial cost	Parameterized model, e.g. in annex B
Grid	Stack at max power when price ≤ 50 €/MWh. Use grid unless local source available.	25000 €	Electricity price time-serie, O&M: 10,000 €/year	Ecoinvent 3.6 high voltage, french mix
Tanks	30-60 bar pressure initial pressure - 30 bar	1135 €/size (m^3)	-	Not included

² https://github.com/romainsacchi/LCA_WIND_DK/blob/master/LCA_parameterized_model_Eolien_public.ipynb

Grid and operation strategy. An operation strategy is necessary for the functioning of energy simulations software used in this thesis. It mainly functions as conflict resolution. In this case, use of electricity from grid or local sources PV/wind. The strategy selected is hydrogen production at maximum power if the price of electricity is less than 50 €/MWh. Since in this particular time serie, there is an excess production from nearby sources, leading to exceptionally low prices during certain time periods. The same strategy was used in the original case study conducted previously in the laboratory for dimensionning with respect to tech-eco. The threshold values between 45-59 €/MWh provide comparable trade-off between the load satisfaction and hydrogen price. One caveat to note is that, when the local sources of electricity, PV and wind are present and available, they are used irrespective of the operation strategy. Production continues as long as the hydrogen can be produced to satisfy the demand or to fill the storage. Impact of the grid electricity is modelled as the high voltage french mix in Ecoinvent 3.6 [Wernet2016].

Hydrogen storage tanks are operated within a range of 30-60 bar. Pressure cannot be dropped below 30 due to technical constraints. The starting pressure is set at a minimum of 30 bar. Since it was observed that when the pressure is set anywhere above 30 bar, the storage was over-dimensionned due the already available 'free' hydrogen in the system. The environmental impact of storage was not included since the goal of the case study was a demonstration of a global approach rather than calculation of exact environmental impact. Ideally, it should indeed be included.

Other LCA inputs. The LCA boundary selected is cradle to grave. It includes as much impacts as possible throughout the project lifetime including raw material extraction, fabrication of components, use phase, transport until their end-of-life. For transport, default values used in the market mixes are used, additional transport in the foreground processes are not included but they can be easily added if required by the users.

The geographical preference for all activities are selected in terms of increasing geographical area around France (e.g. France, neighbouring countries, Europe, Global). These preferences are coded in Brightway2 to select the activities in ecoinvent closest to the desired the geographical scope.

For the end-of-life, only the recycling of materials for which the activities are available in ecoinvent 3.6 (metals & plastics) are included [Wernet2016]. The system is credited for recycling by avoiding the production of virgin materials. The recycling rate is set at 50%. Rest of the materials are either landfilled or incinerated. The calculation steps followed in section 3.3.3 is used again.

End-of-life processes are not included when LCA indicators are calculated in the optimisation loop itself to reduce the computational time. Since, it was verified during preliminary runs that the inclusion of end-of-life processes only marginally reduced the total impact for this case study. As was also seen in chapter 3, the low importance of recycle rate. However, at the end of the optimisation run, when the pareto optimal solutions are finally compiled, impact is calculated including these processes. These processes can indeed be included while maintaining a low computational time. Since the results were not expected to change with this step, it was kept for future enhancements.

5.2.2 Computational structure

Once the integrated model with tech-eco-LCA indicators is ready, the optimisation algorithm can be coupled with it. In essence, the algorithm gives various parameters to the integrated model who then calculates the desired performance indicators or the fitness. The computational structure for this case study is illustrated in in figure 5.3.

NSGA-III first launches energy simulations in Odyssey for different system sizes. The result of this simulation are the tech-eco indicators and the operational details, for each system con-

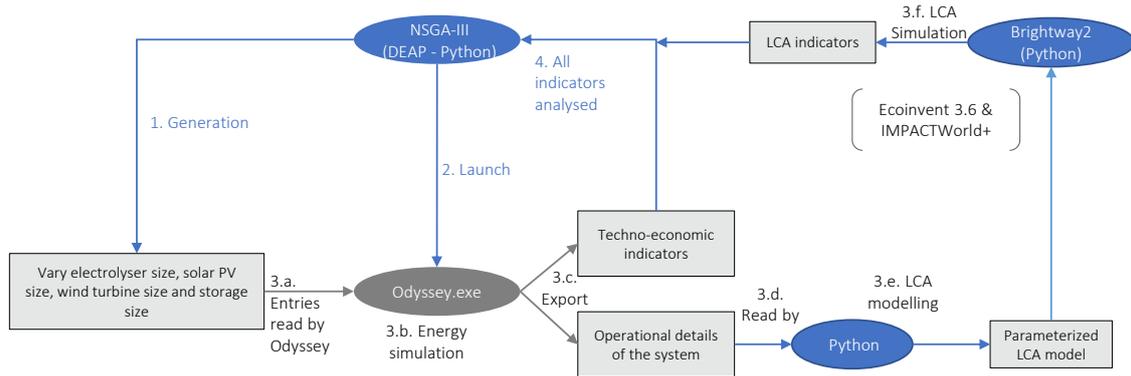


Figure 5.3: Computational structure for search of solutions for including technical, economic and LCA indicators

figuration. Both are read in Python. The operational details are the time-series of performances of each component for one year. This is then used to initialise the parameterized LCA models. The calculation details using them are already specified in section 3.3.1. Additionally for the two new components, the size of wind turbine and solar PV are obtained from energy simulations to initialise their LCA models.

LCA inventories with the different activities is thus compiled followed by impact calculation using Brightway2. All indicators are then analysed by NSGA-III for dimensioning. The loop continues until a stopping criterion is attained.

5.3 Performance of NSGA-III with reduced indicators

In this section, the performance of NSGA-III with reduced set of indicators is tested. Consequently, a comparison between three optimisation cases: one with full set of objectives (14); second with indicators reduced using visual survey (7) and third with indicator reduction using heuristic of [Deb2005, Saxena2012] relying on principal component analysis (2-4). The overall comparison scheme is present in figure 5.4.

From the initial set of indicators (15), the technical indicator, unmet load is added as a constraint at 10%. Individuals below this value were discarded. Since it was observed that often solutions below a certain satisfactory value were termed as technically feasible solutions. Any other value for the 'feasible solutions' could also be defined instead. Or it also possible to include this as an additional objective for optimisation.

The technically feasible system configurations are thus searched for all 14 objectives using NSGA-III. For the second test, 14 objectives are reduced using the visual survey to 7. For the third test, 14 objectives are reduced to 2-4 objectives using PCA [Deb2005]. All LCA impacts for the temporarily discarded indicators are however still retained. Even though they are not used during the optimisation loop, they are used to finally compare the quality of solutions obtained. The three setups are then compared for the no. of function evaluations required for convergence and the quality of solutions obtained. This is done by assessing the pareto optimal solutions, their objective values and the hypervolume indicator. Whereas the objective reduction procedures for the two approaches and reference point selection will be explained later in this section.

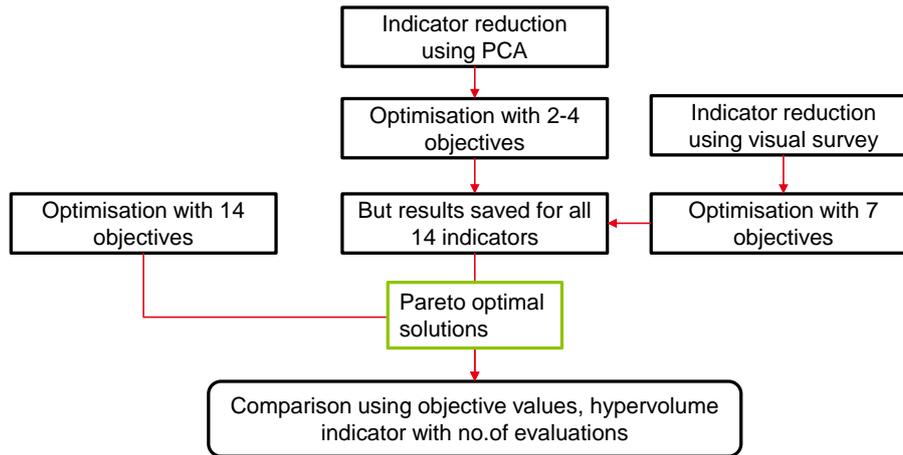


Figure 5.4: Comparison scheme for testing performance for different optimisation cases

5.3.1 Hypervolume indicator

Hypervolume indicator is widely used to measure the quality of solutions proposed by multi-objective optimisation in terms of their convergence and diversity [Bechikh2016]. The former means its closeness to the optimal set while the latter means the spread of solutions in the entire objective space [While2011]. It quantifies this by measuring one single value: the portion of space in all objectives covered by a set of solutions. The space is measured with respect to non-ideal point. This point is generally between 5-50% worse than the worst point expected in the set [Ishibuchi2018]. Hypervolume indicator is especially useful in practical problems when the exact pareto front is not known. It will thus be used in this chapter to assess the quality of the solutions in the pareto front.

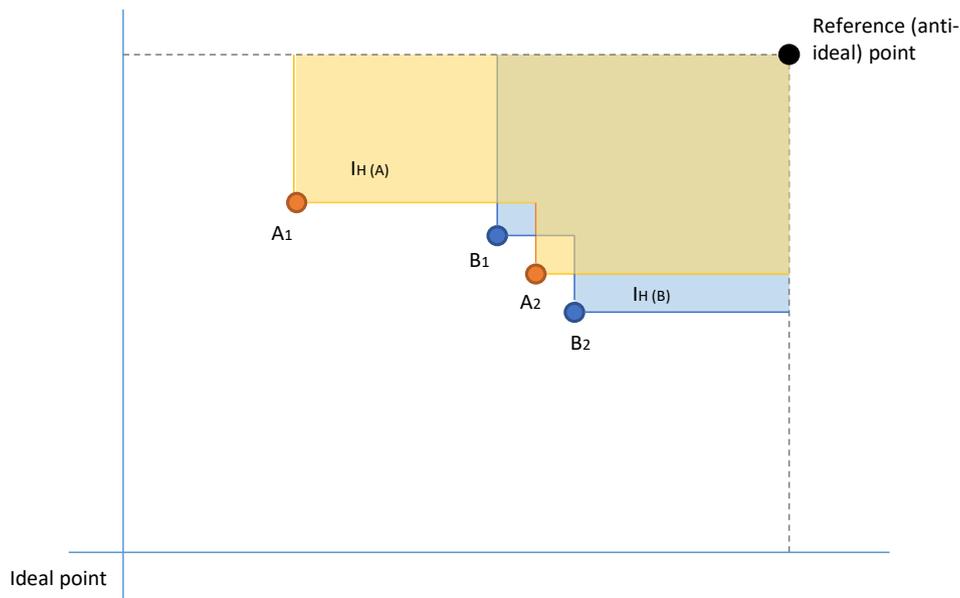


Figure 5.5: Hypervolume calculation principle adapted from [Zitzler2003].

For example, consider two solution sets A and B with two points each. If the goal is to minimise the two objectives f_1 and f_2 , the space covered by the respective solution sets in the two

objective space, $I_{H(A)}$ and $I_{H(B)}$ can be seen in figure 5.5. As the solutions become widely distributed (diversity) or closer to the ideal point (convergence), the space covered by them increases.

Theoretically speaking, four properties of hypervolume indicator found in the literature are important to note:

- If $I_{H(A)} > I_{H(B)}$, there is definitely one space which is dominated by A which is not by B. Thus it must contain a solution vector A_1 which is not dominated by any of the solutions in B [Zitzler2003].
- Whenever a solution set A dominates another set B, then the hypervolume indicator is always greater for the former than the latter [Zitzler2007].
- If a solution set A achieves the maximum possible hypervolume for a given problem, it guarantees that A contains all Pareto-optimal objective vectors [Zitzler2007].
- If A is a preferred alternative to B, it is not possible to have an indicator value as $I_{H(B)} > I_{H(A)}$ [Zitzler2003].

To conclude, if the hypervolume indicator for a given set is higher than the other, it can be concluded that, the given set is at minimum, comparable to the other set in terms of dominance but with higher diversity, like in the above figure. However, it is also possible that the set with higher hypervolume dominates the other set. Thus, solution sets with higher hypervolume values are generally considered of better quality.

The hypervolume calculation algorithm proposed by [While2011], available online³ is used in this thesis. Since the DEAP package uses the hypervolume calculation proposed by [Fonseca2006], for which computational time is infeasible for the number of objectives (14) and size of pareto (1000+) in this thesis.

5.3.2 Indicator reduction using visual survey

The visual survey steps applied to this case study is shown in figure 5.6. The first step is the specification of components in terms of the functional unit 'Impact/kg Hydrogen produced'. Then, an estimation of the impact based on a static simulation of possible impact values attainable by these components are obtained.

In brief, the LCA models described above are utilised again but a simplistic estimation is made based on average values from production and consumption time series. Out of the 4 components, impact of grid and PV will be constant per kg hydrogen produced. Since scaling of their impact is linear with respect to the electricity produced. For PEM electrolyser and wind, their scaling with respect to size is non-linear. Hence, the first step is to find an initial estimate of their sizes based on average values using the following equations:

$$Electrolyser_{size}(kW) = \frac{LHV_{H_2}(kWh/kg) * H_2 \text{ avg. demand } (kg/h)}{efficiency} \quad (5.1)$$

$$Wind_{size}(kW) = \frac{El_{y_{size}}(kW)}{Wind \text{ capacity factor}} \quad (5.2)$$

Where, $LHV_{H_2} = 33.3 kWh/kg$ is the lower heating value of hydrogen. Average hydrogen demand and capacity factor is calculated using the one year time series. The average sizes of electrolyser and wind are 878 kW and 2.05 MW respectively. Then, a range of possible sizes attainable is estimated to be around (500 kW-5 MW) PEM electrolyser and (0-8 MW) for wind

³ <https://pypi.org/project/hvwfg/>

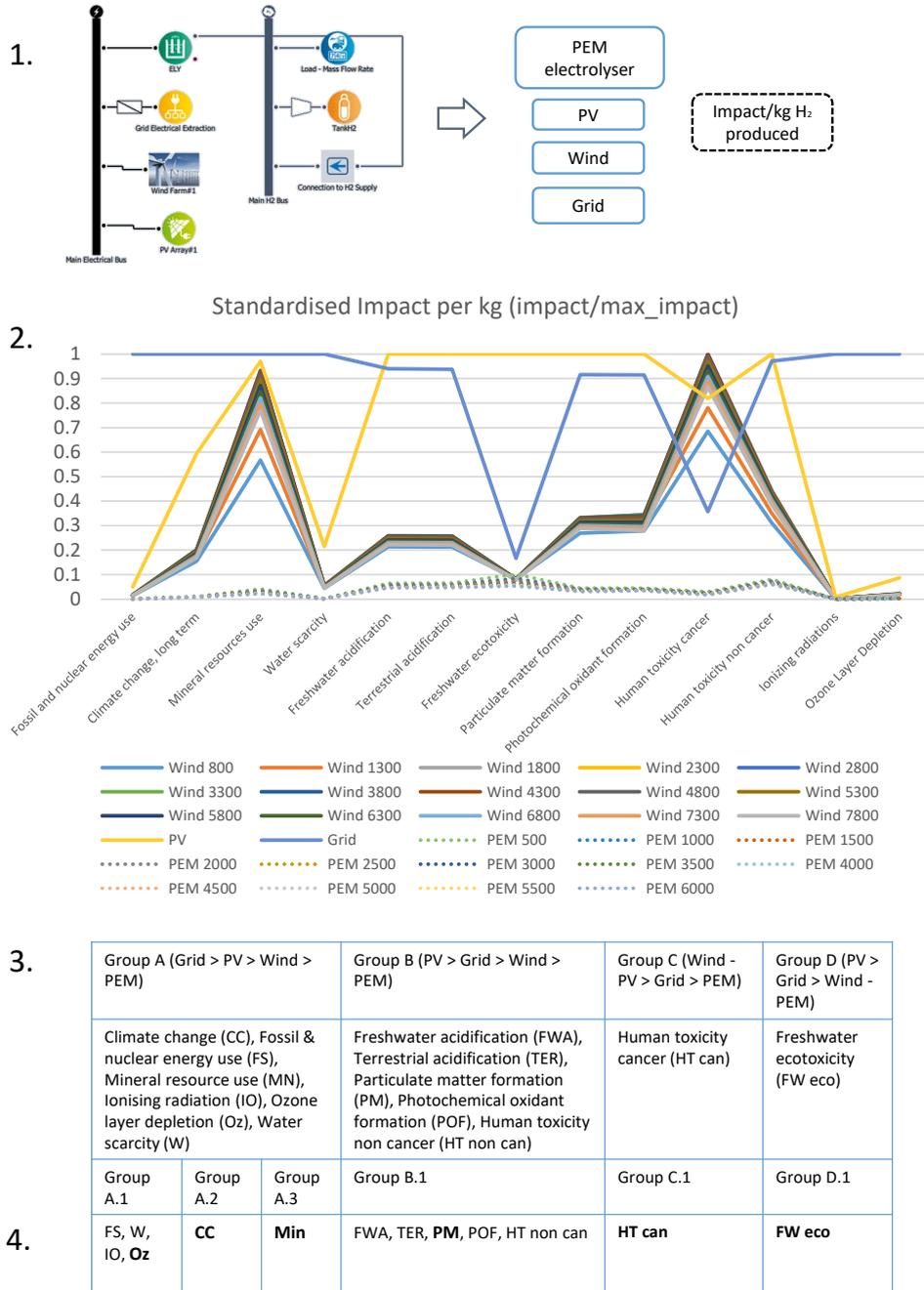


Figure 5.6: Indicator selection procedure using visual survey for this case study

turbine. Equidistant points from this range are selected to calculate their impacts. First for electrolyser, the selected points of sizes, are fed in equation 5.1 to get the average hydrogen produced. Then, the impact per kg hydrogen is calculated using the following equation:

$$Electrolyser(Impact\ per\ kg\ H_2) = \frac{Electrolyser\ cradle\ to\ grave\ impact}{H_2\ avg.\ production\ (kg/h) * 8760h * years\ of\ operation} \quad (5.3)$$

A similar procedure is followed for wind. The equidistant wind sizes in the range (0-8MW) are used first to calculate the electrolyser size from equation 5.2. Then, average hydrogen production from this wind size is calculated using equation 5.1. Finally, equation 5.3 is used to calculate the final impact per kg hydrogen of each wind turbine by substituting the electrolyser impact by wind

LCA models. Similar procedure was also followed for solar PV, but its impact/kg H_2 remains constant for a particular location.

The impact per component is then standardised with respect to the max value in each indicator to obtain the final comparison graph. Ultimately, the indicators are classified into groups and sub-groups. One indicator is then selected at random from each sub-group in bold in the above figure.

5.3.3 Indicator reduction using PCA

Initial 14 indicators were also reduced using the PCA method proposed by [Deb2005,Saxena2012]. It was performed on the pareto front (sample set) obtained after the 5 generations of running NSGA-III with all objectives. This was done for 5 independent runs. It was noted that, a sample set after higher number of generations is generally utilised in the literature [Yuan2017, Saxena2012]. However, due to high computational time (9s) per simulation, from a practical perspective, this limit is selected.

Here, the indicator reduction procedure was varied by changing two settings: the threshold value of including principal components (CUT) and the last step of reducing additional indicators using a correlation matrix. This was done in order to find the influence of these two settings on the final results. Changing the two above steps changes the number and the type of indicators retained for the optimisation. Based on them, insights about these settings for the future use could be obtained. The two instances are described below:

First instance: CUT value = 99.99% and additional reduction using a correlation matrix

Following the steps of schematic 2.10, first a correlation matrix is obtained. Then, the eigenvalues and eigenvectors are calculated. The highest eigenvalue corresponds to the first principal component (PC), the next highest with the second PC and so on. Only the first 5 PCs are displayed in table 5.2 since from 5th PC onwards, they have 0 contribution to the explained variance. As per the CUT = 99.99%, the first four principal components are retained. Then, the objectives with most negative and most positive values are chosen. These are highlighted in bold in table 5.2.

Then, from the 7 selected indicators, additional reductions are made using correlations between them. First, the correlation threshold (T_{cor}) is calculated as per [Saxena2012] using the following equation:

$$T_{cor} = 1.0 - PC_1^{cum} \left(1.0 - \frac{N_{pc}^{cum=95.4}}{N_{OBJ}} \right) \quad (5.4)$$

Where,

PC_1^{cum} : is the cumulative variance explained by the first principal component

$N_{pc}^{cum=95.4}$: Number of principal components required to explain 95.4% of the variance

N_{OBJ} : Number of objectives in the original problem

Using the values from the table 5.2, $T_{cor}=0.4$. The 7 indicators obtained are classified using the correlation matrix such that their correlations coefficients are greater than the calculated T_{cor} . The correlation matrix is included in annex table A.12. Consequently, from this step, two sets of objectives are formed. The first one being water scarcity and ionising radiation. The second set includes rest of the 5 objectives. Then, one objective from each set is selected which has the highest score (sc_i) defined by [Saxena2012]:

$$sc_i = \sum_{j=1}^{N_{pc}} PC_j \cdot abs(ev_{ij}) \quad (5.5)$$

Table 5.2: PCA results for the 14 indicators. Selected indicators are highlighted in bold.

	PC ₁	PC ₂	PC ₃	PC ₄	PC ₅
Eigenvalues (λ)	9.82	3.77	0.36	0.05	0.00
Explained variability (%)	70.11	27.93	2.6	0.35	0.0
Cumulative explained var. (%)	70.11	97.04	99.65	99.99	99.99
Cost in €/kg	0.282	-0.109	-0.654	-0.649	0.008
Fossil and Nuclear energy (FS)	-0.206	0.388	-0.218	0.016	0.074
Climate change (CC)	0.249	0.320	0.013	0.226	-0.241
Mineral resources (Min)	0.303	-0.096	-0.379	0.469	-0.211
Water scarcity (W)	-0.099	0.489	-0.087	0.051	-0.013
Freshwater acidification (FWA)	0.292	0.203	0.139	-0.089	-0.176
Terrestrial acidification (TER)	0.292	0.204	0.142	-0.093	-0.174
Freshwater ecotoxicity (FW eco)	0.297	0.138	0.398	-0.286	-0.131
Particular matter formation (PM)	0.306	0.146	0.054	0.090	0.902
Photochemical oxidant formation (POF)	0.307	0.136	0.076	0.117	-0.004
Human toxicity cancer (HT can)	0.305	-0.112	-0.293	0.389	0.016
Human toxicity non cancer (HT non can)	0.298	0.182	0.030	-0.181	-0.013
Ionizing radiation (IO)	-0.224	0.361	-0.220	-0.004	-0.002
Ozone layer depletion (Oz)	-0.189	0.412	-0.178	-0.009	-0.004

N_{pc} : total number of principal components

PC_j : explained variance of the principal component 'j'

ev_{ij} : is the i^{th} eigenvector of the j^{th} PC

Thus for example from table 5.2, $sc_{cost} = 0.282 * 0.7 + 3.77 * 0.109 \dots$ and so on. In this way, the indicators having the maximum contribution to all the PC out of the reduced set are selected. This results in selection of 2 indicators: *freshwater ecotoxicity* and *ionising radiations*. This procedure was repeated 3 times, it resulted in the selection of the same 2 indicators.

Second instance: CUT value = 95% and no additional reduction

In the second instance of using PCA, the last step is skipped which involves additional reduction performed using correlation matrix. The rationale behind this was that, it was noticed during the reduction to 2 objectives in the previous step gave comparatively worse final pareto solutions, outlined by a low hypervolume indicator and distribution of points. These topics will be discussed in the results.

If CUT value was maintained at 99.99%, without additional reduction in indicators, 7 indicators were obtained, that would be very similar to the ones selected by the visual survey which will most likely lead to similar performance. A comparison is seen in table 5.6. By reducing the CUT value and no additional reduction, 3-4 indicators could be selected that could offer insights on the trade-off between number of indicators and solution quality.

Thus PCA reduction procedure is tested to include only the first two principal components using a CUT value of 95%. This value was also used in the other investigations in the literature

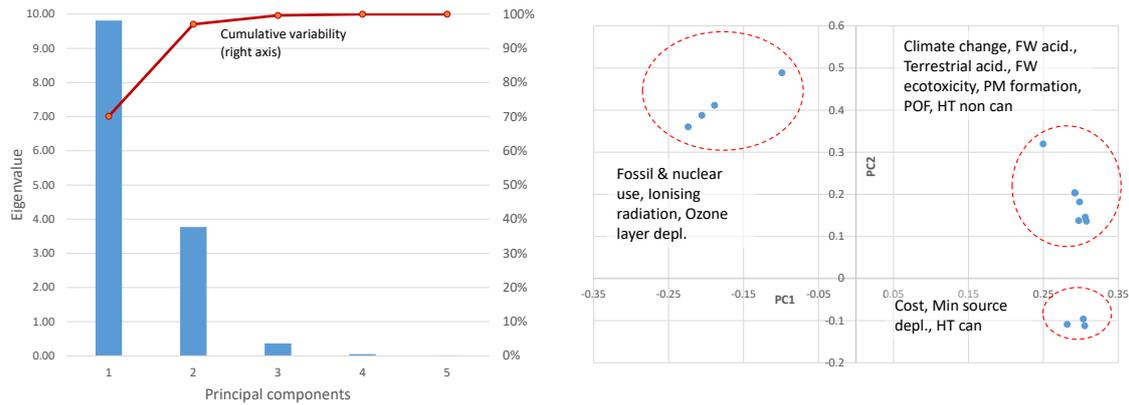


Figure 5.7: (Left) Scree plot of principal components are their cumulative explained variability. (Right) First two principal components plotted against each other to reveal correlations between indicators.

[Perez Gallardo2013, Pozo2012] according to the first proposal of the reduction framework in [Deb2005]. Thus from table 5.2, only indicators with the most positive and most negative values are selected. This is done for five simulation runs, the indicators retained for 5 runs are presented in table 5.3. For the 3rd simulation run, human toxicity cancer was selected twice in the two PCs. Hence, only 3 indicators were retained.

The scree plot of the PCs from table 5.2 is illustrated in on the left in figure 5.7. On the right, the two PCs obtained are plotted against each other where indicators with similar trends are grouped together.

Table 5.3: Indicators retained by PCA during the five simulation runs

Simulation run	1	2	3	4	5
Indicators Retained	Ionising radiation	Ionising radiation	Ionising radiation	Ionising radiation	Ionising radiation
	Photochemical oxidant formation	Photochemical oxidant formation	Human toxicity cancer	Photochemical oxidant formation	Photochemical oxidant formation
	Water scarcity	Water scarcity	Water scarcity	Water scarcity	Climate change
	Human toxicity cancer	Human toxicity cancer		Human toxicity cancer	Human toxicity cancer

5.3.4 Reference point specification

Reference point and consequently population size selection is important for analysing the performance of the 3 instances. Since it will directly influence the no. of function evaluations for each generation and thus the time required for each simulation run. As seen in equation 2.12, since no. of objectives (N_{OBJ}) are fixed, axis partitions (\hat{P}) are needed for defining the number of uniformly distributed reference points. The higher the number of partitions, the better it is since more diverse solutions for each objective can be obtained. This number then decides the number of reference points and consequently the population size. Following the authors, population size is slightly higher than the reference points [Deb2013].

Table 5.4: Reference points, population size and indicators retained for the different simulation runs.

Instance	Original objectives	Reduction using visual survey	Reduction using PCA		Reduction using PCA (2 nd instance)
			(1 st instance)		
No. of objectives (N_{OBJ})	14	7	4	3	2
Axis partitions (\hat{P})	2,2	3,2	6	11	72
Layer scale	1,0.5	1,0.5	1	1	1
No. of reference points (H)	210	112	84	78	73
Population size	212	120	92	84	80
Indicators for search	Cost, FS, W, IO, Oz, CC, Min, FWA, TER, PM, POF, HT non can, HT can, FW eco	Cost, Oz, CC, Min, PM, HT can, FW eco	IO, HT can, POF, W/CC	IO, W, HT can	IO, Fw eco

The population size is also important, since it should be large enough to ensure that sufficiently different individuals are present in generation zero. Thus first a minimum population size is decided to be 80.

For 2 objectives, 72 axis partitions result only in 73 reference points and minimum population size. Then, 3 objectives, 11 axis partitions can be afforded which result in 78 reference points. Whereas the axis partitions are reduced for 4 objectives to get 84 reference points. From 7 objective onwards, layering is needed. Since even with $\hat{P}=5$, H becomes 462. Thus 2 layers, one outer layer with 3 axis partitions and second inner layer with 2 axis partitions are selected. For 14 objectives, \hat{P} for outer layer has to be reduced to 2, otherwise $H=665$. It cannot be reduced further since as seen in figure 2.7, for $\hat{P}=1$ only the extreme points of each objective are selected. An overview for the different cases is provided in table 5.4.

5.3.5 Results with NSGA-III and reduced indicators

Hypervolume indicator for the different number of reduced objectives with NSGA-III are shown in figure 5.8. As mentioned before, hypervolume indicator signifies convergence as well as diversity of solutions. Higher the indicator value, solutions are considered of better quality. It is calculated for the pareto solutions obtained after each generation for each simulation run. As outlined before, the hypervolume indicator was calculated with respect to all 14 objectives irrespective of when the objectives were temporarily omitted during the search using NSGA-III as seen in figure 5.4. The anti-ideal point for the indicator calculation was kept the same, which corresponds to approximately 30% worse objective values than the maximum obtained in the preliminary simulation runs.

It can be seen that for all five runs, optimising with the reduced 7 objectives has a higher hypervolume than the original 14 objectives. The latter case struggles to reach a similar value even for the same amount generations or more than twice the number of evaluations. This can be verified from annex figures C.1 and C.2.

On the other hand using PCA, the indicators were reduced to between 2-4 indicators depending on the instance. For the first instance, a 99.99% CUT value and following all steps recommended

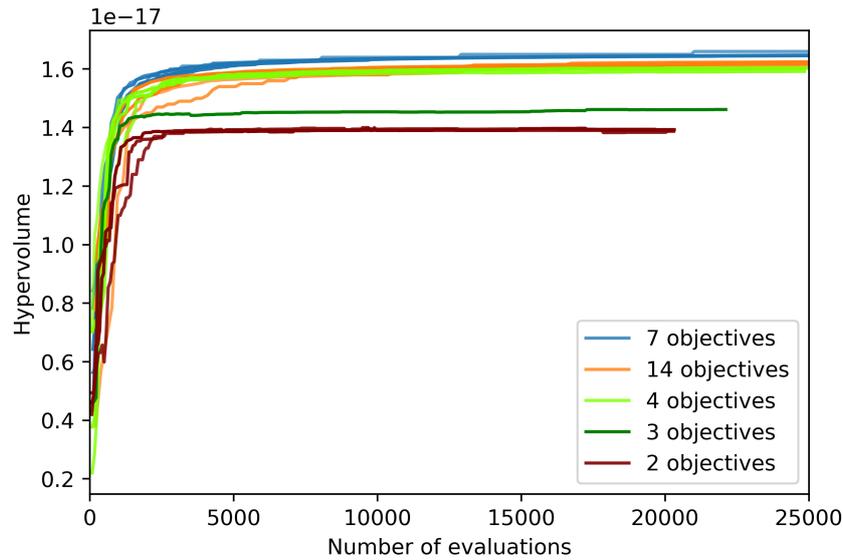


Figure 5.8: Hypervolume indicator for different simulation runs.

by [Saxena2012], 2 objectives were retained. Whereas for second instance, with 95% CUT value and skipping additional reduction, 3/4 indicators were obtained. For all of the reductions using the above settings in PCA, hypervolume indicator was the less than with 7 or 14 objective cases. In fact, as the indicators were reduced beyond 7, hypervolume also similarly reduced.

It should be noted that according to the specified reference points and consequently the population, the number of evaluations for each generation differs depending on the number of objectives optimised. For example, optimisation with 14 objectives requires approximately 2 times more function evaluations than with 7 objectives due to twice the population size. Hence, comparison of hypervolume is fair with respect to the number of evaluations. In any case, the above observations do not change even when a comparison is made with respect to the number of generations.

2D plots of pareto front In order to have a closer look at the pareto solutions provided by the different simulation runs, they are plotted on four bi-dimensional graphs. These are all pareto optimal solutions in 'n' objective space. Where 'n' depends on the no. of objectives (14, 7, 4..), can be seen in the graph legends. These pareto solutions were plotted with respect to all 14 objectives in bi-dimensional plots at the end of 300 generations. The number of individuals in the pareto here are 9951, 8781 and 3615 for 14, 7 and 4 objectives cases plotted in this section respectively.

In figure 5.9, pareto solutions obtained by one simulation run of 14 and 7 objectives can be seen. While in figure 5.10, comparison is made between 4 and 7 objective runs. An optimisation with respect to two objective will give a single curve of points as was seen in figure 1.11. Here since the solutions optimal in higher dimensions are plotted in two dimensions, the pareto curves are thicker. Same reason also explains the different shapes of these curves as compared to a 2 dimensional pareto curve.

In the plots of climate change vs. freshwater ecotoxicity, discrete points are observed, this is because of the discontinuous search space provided. For example, for wind turbine sizing, system configurations were searched with steps of 100kW between 0-8MW.

In both figures the superior distribution of solutions obtained with 7 objectives can be seen. Hence, validating also the higher hypervolume indicator obtained for this instance. The worse distribution obtained with 14 objectives is likely due to the less distributed reference points possible. In order to limit the size of population and consequently the number of evaluations, a layering approach was used to specify reference points as per the authors [Deb2013]. Population for each

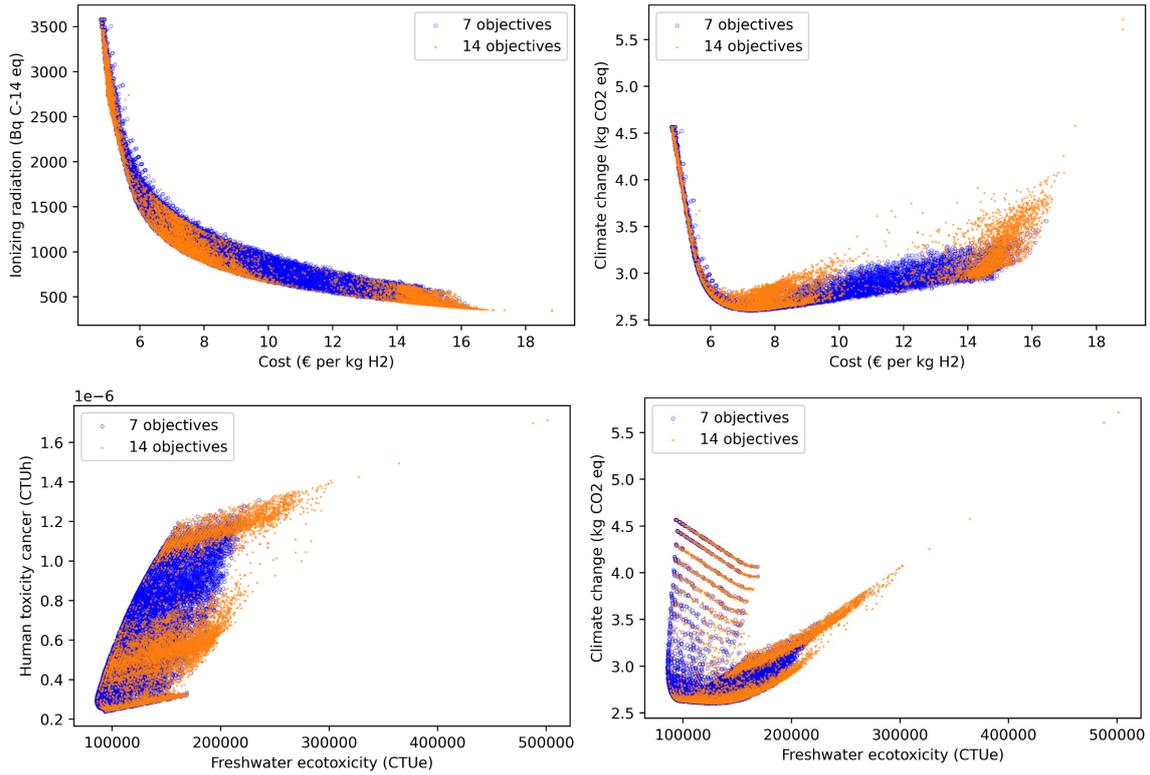


Figure 5.9: Distribution of the pareto front obtained using 14 and 7 objectives shown in four figures of 2-dimensional plots

generation is selected after assigning each individual to a reference point. Then, the pareto front is updated using this population if it is non-dominated compared to the individuals already present. Thus it is likely that due to insufficiently distributed reference points, spread of solutions obtained also follow a similar trend. It should be noted that the optimisation with 14 objectives has higher number of evaluations and higher number of pareto optimal points. Irrespective of this, it struggles to uniformly distribute the points, as can be seen in figure 5.9. Despite having high number of points, they leave a noticeable empty space in the plots.

For optimisation with 7 objectives, despite the lower number of pareto optimal points, they seem to be better distributed. Here a layering approach was also used for reference point distribution. The number of reference points are decreased to half as compared to the original search space, but they are better distributed in the reduced 7 objective space. Thus, better distributed solutions in the reduced space could be found. Furthermore, since these solutions were highly correlated with rest of 7 temporarily omitted objectives, the search provides a better distribution of solutions in all objectives.

On the other hand, if we look at the solutions with 4 objectives, they have a high distribution of reference points but the same does not reflect in the pareto solutions obtained. In cost curves (top two figures in 5.10), insufficient distribution is found towards the cost minima. Whereas in the bottom two graphs, there seems to be a significant portion left unexplored by a 4 objective search. The 4 objectives correspond to two PC retention, in the PCA reduction procedure that explain 97% of the cumulative variance. As a result, 3% of the variance of the original objective space is not explained in this instance. Thus solutions optimal in the higher dimensional space might have been discarded when optimising with 4 objectives. The inferior performance is likely due to this reason.

The optimisation instance with 2 reduced objectives obtained from PCA, resulted in even worse performance. This could be expected also due to its low hypervolume indicator value ob-

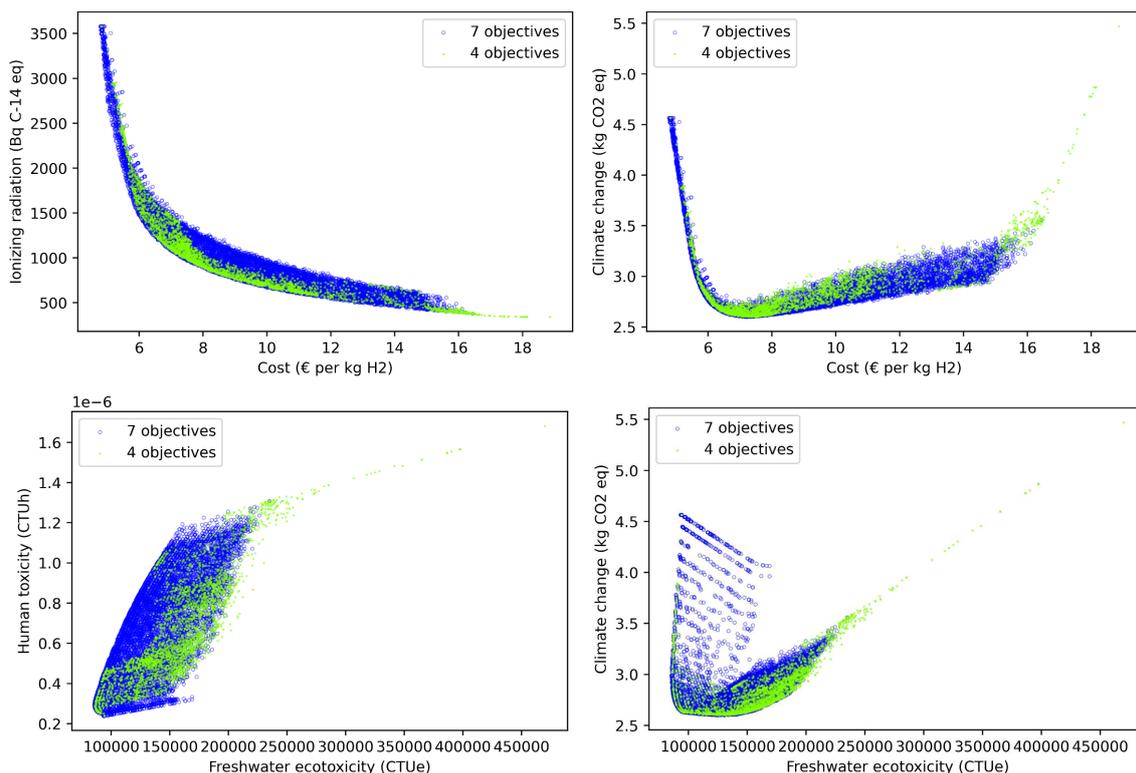


Figure 5.10: Distribution of the pareto front obtained using 7 and 4 objectives shown in four figures of 2-dimensional plots

tained. The 2D plots for two simulations for 2 and 7 objectives can be found in annex figures C.7 and C.8.

There are although one set of points which are comparatively less explored by the 7 objective optimisation, which are included in the other two, seen in the above figures, for costs higher than 16 €/kg. These points correspond to extreme minimisation of ionising radiation, ozone layer depletion, water scarcity and fossil & nuclear energy use. As a result, wind and PV are oversized in an attempt to minimise the grid electricity use. However, climate change increases instead of decreasing since oversized wind and PV lead to much of the electricity that is wasted and not used by the electrolyser to produce hydrogen. This also results in higher cost per kg hydrogen and high impact per kg in all other indicators.

To explain why these points are not included in 7 objective optimisation, see two plots in figure 5.11. In the reduction using visual survey, ozone layer depletion was selected, which is highly correlated with ionising radiations. Even though the overall correlation coefficient equals 1, the perfect correlation does not hold true at the extremities during the minimisation of ozone layer depletion and ionising radiation (see figure 5.11 on the right). Thus, when ozone layer depletion is minimised less than 0.7 kg CFC-11 eq ($\times 10^{-6}$), the points are not optimal in the reduced 7 dimensional plane anymore. Points above 16 €/kg H₂, in the figure 5.11 on the left. Whereas since ionising radiation is included in the reduced set proposed by PCA, these solutions are identified in this instance. This can be seen as a limitation of correlation based objective reduction techniques. Sometimes the extreme points might not conform to the same correlations.

All of the observations seen above regarding the distribution of points are the same for different optimisation instances. Same graphs for two other independent runs are present in the appendix, they are almost indistinguishable with the ones shown above. Plots for 7-14 objectives can be found in annex figures C.3 and C.4. While for 7-4 objectives, can be found in figures C.5 and C.6.

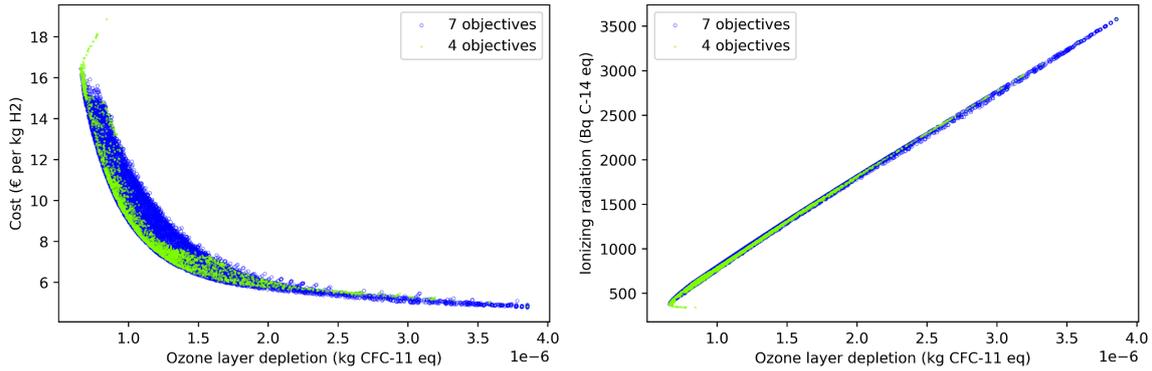


Figure 5.11: Distribution of the Pareto front obtained using 7 and 4 objectives shown for two, 2D plots

Evaluations to reach minimum value of indicators In terms of searching the minimum value of each objective, these values were found relatively quickly for all instances with (7 and 4) objectives being slightly faster. The minimum value of indicators in the Pareto were about the same for all simulation runs except for the 2 objective case.

A standardised minimum value of Pareto solutions for four objectives plotted with respect to the number of evaluations are illustrated in figure 5.12. Since the performance of the 2 objective case was considerably worse and to ease the visualization it is not plotted here. But the same figures including it can be found in annex figure C.9. A comparatively lower value for some indicators such as cost was not reached even after 300 generations for this case.

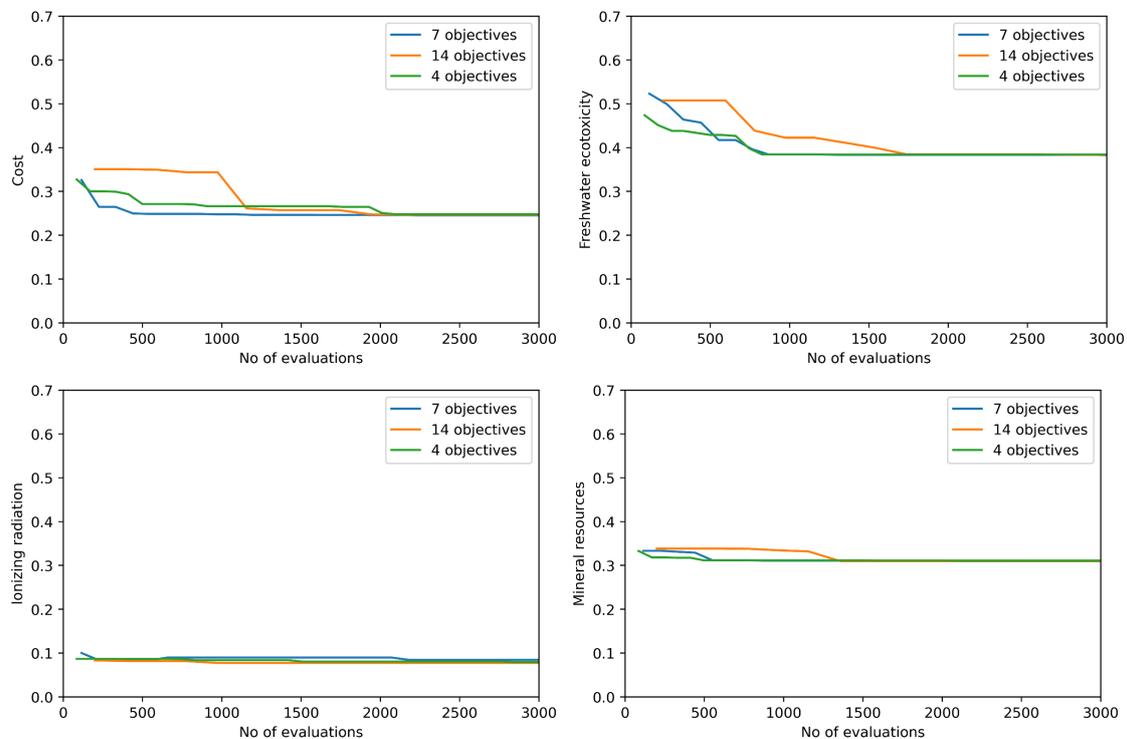


Figure 5.12: Standardised minimum value of objectives vs. number of evaluations for different simulation runs.

The minimum value of the objectives are standardised by dividing with the anti-ideal point used for hypervolume. Such standardisation is important since LCA indicators show a difference in magnitude of up to 10^{12} between them. Thus, a small difference in indicators with high magnitude may seem significant and vice-versa. In most indicators such as cost, freshwater ecotoxicity and mineral resource depletion, less objectives lead to relatively faster search of the minimum objective value. While in others such as ionizing radiations, no difference is found. However, it should be noted that finding minimum points might not reflect optimal solutions or a satisfactory distribution of pareto points for decision-making. For all simulation runs, there was a negligible change in the hypervolume and the indicators after around 200 generations, hence being conservative, the simulations were run for a maximum of 300 generations each.

5.3.6 Discussion

Performance of NSGA-III

It was found that NSGA-III struggles to provide a uniform distribution of solutions with 14 objectives. Even though its capabilities has been demonstrated in other investigations in the literature with equivalent or higher number of objectives. It goes to show that the performance of algorithms are often problem dependent, which was also highlighted by [Li2018]. Nevertheless, NSGA-III performed well with most problems and no one algorithm has the best performance with all types of problems (ibid.). The issue here is likely due to the inadequate distribution of the reference points. The number of reference points could be increased but it will lead to exponential increase in computational time. For example, increasing axis partitions from [2,2] to [3,2] in table 5.4 leads to 665 reference points for 14 objectives. This will result in 3 times more computational time.

Performance of NSGA-III with reduced objectives

The best quality of solutions were found when NSGA-III was run with the 7 reduced objectives. This was verified using the hypervolume indicator values and distribution of pareto optimal points. The probable reason behind this is for the case of 7 objectives, the best trade-off was obtained between: a good distribution of the supplied reference points to NSGA-III and enough indicators retained to explain the variance of the original 14 objective problem.

If the indicators were reduced further, a corresponding decrease in the quality of solutions was noticed. For optimisation with less than 7 indicators, a lower hypervolume indicator was obtained than for the ones found with the original 14 objective optimisation. The worst solution quality was obtained when problem was simplified to 2 objectives.

Time comparison

As mentioned before, the hybrid energy simulations are computationally expensive. One simulation takes around 9 seconds to complete on one core. Out of it, majority of the time is used by the energy simulation model ($\approx 8s$) while the LCA model on Brightway2 takes the rest ($\approx 1s$). The latter can be further reduced using aggregated pre-samples⁴ [Lesage2018]. Estimation of the time required is made in table 5.5 using the population size calculated in table 5.4, as the approximate number of evaluations in each generation. The time required to generate samples at 5 generations of 14 objectives is not included for reduced objectives. Quality of the obtained solutions above are qualitatively rated.

Indicators when reduced to 7 in the case study give the best quality solutions in lesser time. If indicators are reduced further, a trade-off between the time required and the solution quality could be seen. The time required can be further reduced using more processing power by utilising other

⁴ <https://bw2preagg.readthedocs.io/en/latest/>

Table 5.5: Estimation of computational time using multi-processing for different number of indicators.

	NSGA-III with 14 indicators	NSGA-III with reduction to 7 indicators	NSGA-III with reduction to 4 indicators	NSGA-III with reduction to 2 indicators
Time required for 200 generations on 8 cores (Intel i7, 7 th gen)	13 hours	7.5 hours	5.8 hours	5 hours
Quality of solutions	++	+++	++	+

computers on the network. This is often done in the laboratory. Another advantage of genetic algorithms can be seen here. Large problems could be processed in parallel.

Objective reduction using PCA

The first observation is that LCA can indeed be considered as a problem with high number of redundant indicators, meaning that the number of reduced objectives explaining most of the variance are much smaller than the original objective set. This was also observed in other investigations [Steinmann2016, Perez Gallardo2013, Sabio2012].

In the update of objective reduction framework using PCA [Saxena2012], a low CUT value and additional reduction using a correlation matrix are recommended for problems with high number of redundant indicators. When these steps were followed in section 5.3.3, only 2 indicators were retained. Out of all the simulation runs, worst quality of solutions were obtained in this case. Considerably better solution quality in terms of convergence as well as their diversity was obtained with 3, 4 and 7 objectives.

While it might be true that 2 indicators explain most of the variance, a small amount of variance left out in the indicators resulted in much worse solutions by significantly distorting the original pareto structure.

Objective reduction using visual survey

Objective reduction using only a visual survey offered the best solution quality in combination with NSGA-III. The correlations between the indicators were successfully identified even before initialising the problem. Whereas PCA based reductions were carried out after running the 14 objective problem for 5 generations. This can save significant amount of time. Furthermore for this sample set, much higher number of generations are used in the literature, e.g. 200 generations used in [Yuan2017]. Which is not certainly not feasible for application here.

Similar objectives can also be obtained using PCA reduction procedure as well. For example, 7 indicators are also selected if these changes in the PCA procedure are made: the CUT value in the PCA reduction procedure increase to 99.99% and no additional reduction. Additionally, they are almost the same as one selected by the visual survey. A comparison can be seen in table 5.6 below.

If PCA reduction is made using the recommended settings, 4 indicators are the same as obtained for visual technique. While the other three are highly correlated. It is thus highly likely that the quality solutions obtained after optimisation using them are similar to the ones obtained using the 7 indicators in this section.

Table 5.6: Comparison of indicators selected using the visual technique and using a particular setting in PCA.

Indicators selected using visual technique	Oz	PM	CC	HT can	Cost	Fw eco	Min
Indicators selected using PCA (CUT=99.99% and no additional reduction using correlation matrix)	IO	POF	W	HT can	Cost	Fw eco	Min

5.4 Visualization of solutions for decision-making

In this section, an approach for visualisation of solutions from the obtained pareto distribution is presented. As mentioned earlier, there exists other approaches for the same purpose such as multi-criteria decision making [Perez Gallardo2013], data envelopment analysis [Gonzalez-Garay2018], pareto filters [Antipova2015]. A simple alternative approach is presented for the selection that takes advantage of the NSGA-III operation + correlations based objective reduction. Final results are presented in form of a visual tool to the users.

5.4.1 Proposed procedure

Once the optimisation loop is stopped, two groups of solutions are obtained. The population from the final generation and the pareto front, this can be verified from algorithm operation in figure 2.6. A final set of solutions for decision making are to be assembled from these two groups. A schematic of the recommended procedure to obtain these solutions is presented in figure 5.13. In order to select a set of solutions for final decision making, the latest population obtained in the last generation is the recommended starting point since it represents the preferences that were specified before start of the search process.

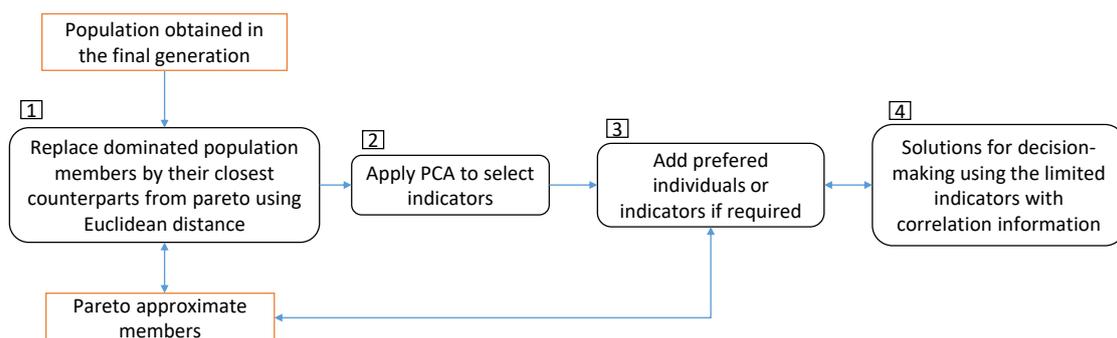


Figure 5.13: Proposed approach to select and visualize solutions.

As seen in the algorithm schematic in figure 2.6, the pareto front is updated using the population obtained at each generation. Thus, it is not necessary that all of the population members are the part of the pareto front. In other words, there could already be members in the pareto front who have better fitness values. In such a situation, the dominated members are replaced by their closest counterparts present in the pareto front in *step 1*. Then to reduce the visual burden, PCA is performed again in *step 2*. Reduced set of indicators are thus obtained that explain most of

the variance. Here the principal components retained can be selected by the users themselves. A threshold value of 95-99% is recommended.

Then, preferred indicators or individuals can be further added in *step 3*. Final results are presented with these reduced indicators with their correlation information with other indicators. Engineers or decision-makers can thus select a solution according to their preferences using a visual tool in *step 4*. If required, preferred indicators can be added to the final visualisation based on correlations. *Step 3 and 4* are iterative where additional solutions could be added to the visual tool for decision-making. These steps are described and demonstrated below.

5.4.2 Application to the obtained results

In this case study, a uniform set of reference points were specified irrespective of the number of objectives. Thus, the final population should represent equidistant spread of solutions on or near the pareto front for all objectives.

Selecting closest counterparts using Euclidean distance. For the demonstration of this subsection, one of the optimisation runs with 7 objectives are selected. Since best quality of solutions were obtained in these runs. At the end of 300 generations, the size of the pareto front is 8781 solutions. We start with 120 solutions in the population, out of which, 37 solutions were found not to be pareto-optimal. Since they are dominated by at least one solution in the pareto front. A particular solution is a system configuration with unique sizes of electrolyser, wind turbine, PV and hydrogen storage. These solutions were then replaced by their closest counter parts in the pareto set. For this, Euclidean distance ' d_{ec} ' between each of 37 solutions and the each pareto set point was calculated using equation 5.6. Then the non-optimal point was replaced by the one with the minimum distance in the pareto set.

$$d_{ec}(m, n) = \sqrt{\sum_{o=1}^z (m_o - n_o)^2} \quad (5.6)$$

Where:

n : is the non-pareto optimal population member

m : is the member of pareto front

m_o, n_o : vectors associated with members m and n respectively of length ' z '

The vectors m_o, n_o used for calculating the distance could either be the standardised fitness values or the system configurations themselves. For the former, fitness has to be divided with the max fitness value for fair calculation across all indicators. Both of these approaches were tried, sometimes they led to same solutions and sometimes not. However, neither led to a better performance than the other. Similarities in the system configurations was selected here, but the other approach could also be selected instead. An example is provided in figure 5.14. Whereas the distribution of the population, after the replacement in terms of ionising radiation and cost is shown in figure 5.15. Similar uniform spread is also observed for other indicators. Thus validating the claim that the final population represents a uniform spread of solutions. That is it represents the preferential solutions defined at the beginning as the reference points of NSGA-III.

PCA for identifying correlations. At this stage, PCA is used to find correlations between the indicators so as to reduce visual burden of having to make a decision based on many indicators. This reduction might have already performed in one of the previous steps. Nonetheless, PCA reduction step is not computationally expensive and easy to implement. Furthermore, PCA at this step can give accurate correlations of indicators amongst the selected pareto solutions. As

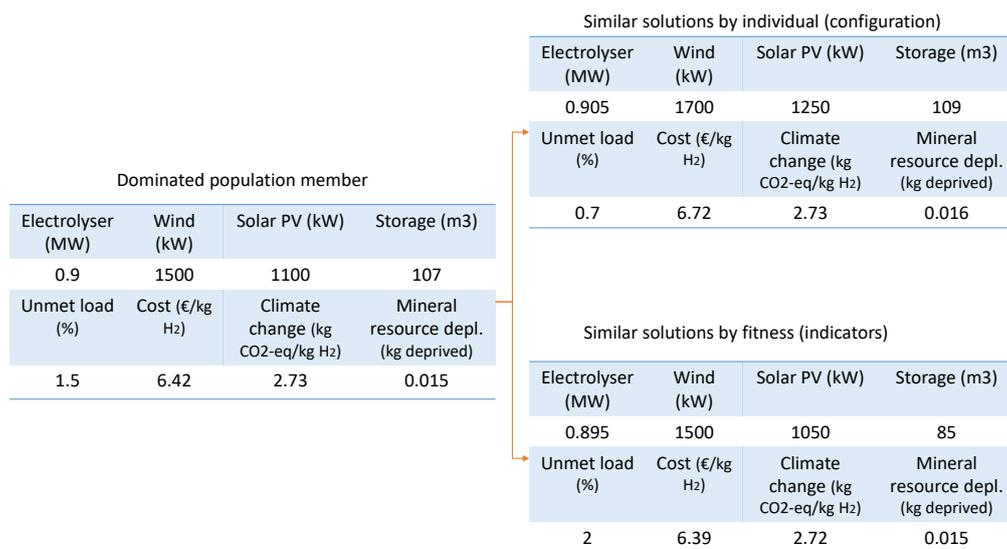


Figure 5.14: Replacing the dominated population member by a similar individual in the pareto front using individual or fitness similarities.

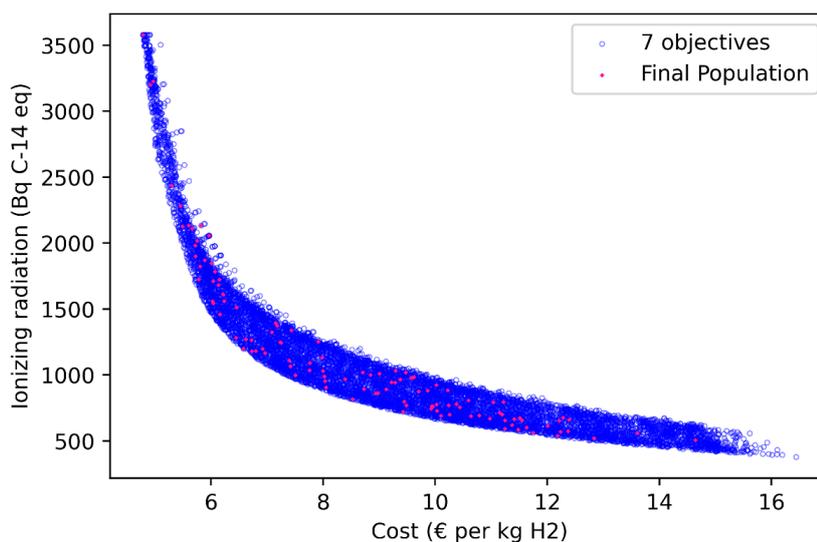


Figure 5.15: Population spread in the pareto.

compared to the optimisation step, more leeway could be afforded for the CUT value, but it is recommended to kept above 95%.

Therefore, for the above case PCA is conducted on the 120 optimal and technically feasible solutions. Three principal components are retained that explain 99.93% of the variance. Four indicators out of the 14 initial ones, are given by this step: Human toxicity cancer, Ionising radiation, Terrestrial acidification and Freshwater ecotoxicity. If required, these indicators can also be replaced by the ones considered more important. In the laboratory, cost and climate change impact are used regularly in most of the projects. Thus they are added, replacing the ones with whom they show the highest correlations: human toxicity cancer and ionising radiation respectively. The correlation matrix for this step is added in the annex table A.13. A change in correlation values could be seen from the previously calculated correlation matrix after 5 generations, which is normal since those solutions were not pareto optimal. Correlations between indicators for solutions on the pareto might be different than the ones that are not optimal [Yuan2017].

Visualisation of solutions with limited indicators Solutions can then be visualized for decision-making. Along with the selected indicators, it is crucial to provide correlation information to the users. That is, information about change in an indicator will lead to corresponding change in which other indicators in the background.

For this case, the solutions were sorted in terms of increasing cost indicator. They can then be selected by users based on the particular project preferences using a visual hovering tool. For the moment, absolute values tech-eco and 2 LCA indicators are shown along with information about each system configuration when a pointer is hovered on it. A demonstration of first 25 solutions is provided in figure 5.16. It should be noted that users can add other indicators to display or even additional performance indicators such as capital cost can also be added easily in the future. The display here is limited but additional solutions can be viewed on a computer screen with or without a scroll bar. Additional solutions could be added by the users from the bi-dimensional plots; or by searching the pareto front using filters such as objective values, system configurations; or by finding similar solutions using euclidean distance; other information such as fixed cost could also added in the future as required.

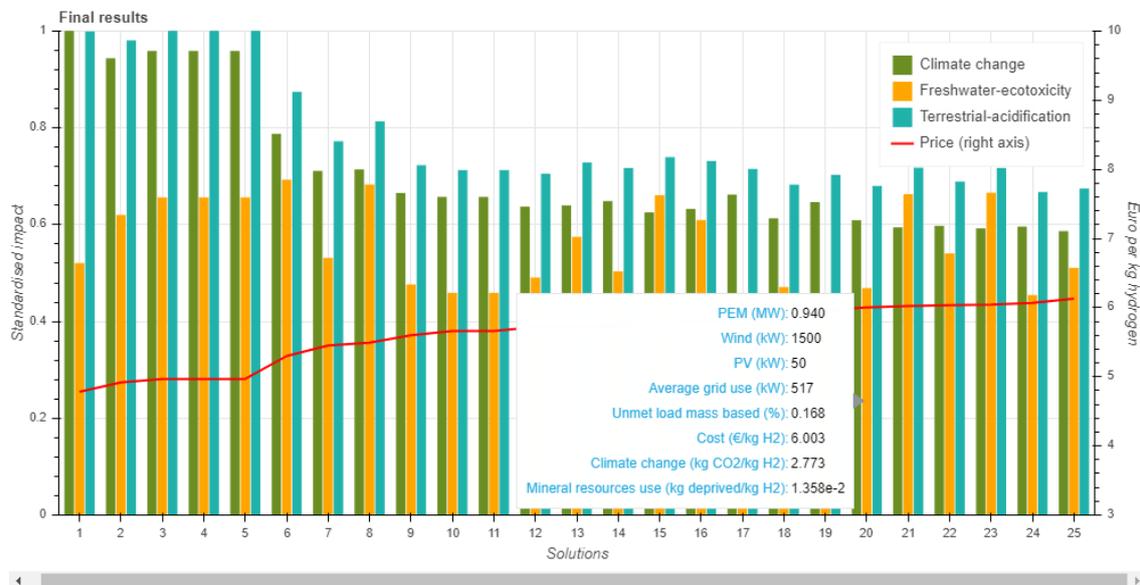


Figure 5.16: Final set of solutions provided in a visual tool.

The indicators which were not selected in the visual tool in the previous step are presented in table 5.7 with the selected indicators with whom they have the highest coefficient of correlation. Similarly, a graphical illustration of this correlation is in figure 5.17. The standardised impacts of the LCA indicators are shown for all 120 solutions. It can be clearly seen that indicators under each group are highly correlated. As per [Taylor1990], all coefficients lie between high (0.68-0.9) to very high (0.9-1) correlations.

In most cases, decision-making based on a small set will work as expected, as can be verified from the plots. However, in certain cases especially at the ends, caution has to be exercised by the users. For instance, in group A in figure 5.17, an increase in cost between solutions 1-5 leads to a decrease in human toxicity non cancer and mineral resource depletion. Thus providing these graphs and correlation coefficients along with the final set of solutions is important for consistent decision-making.

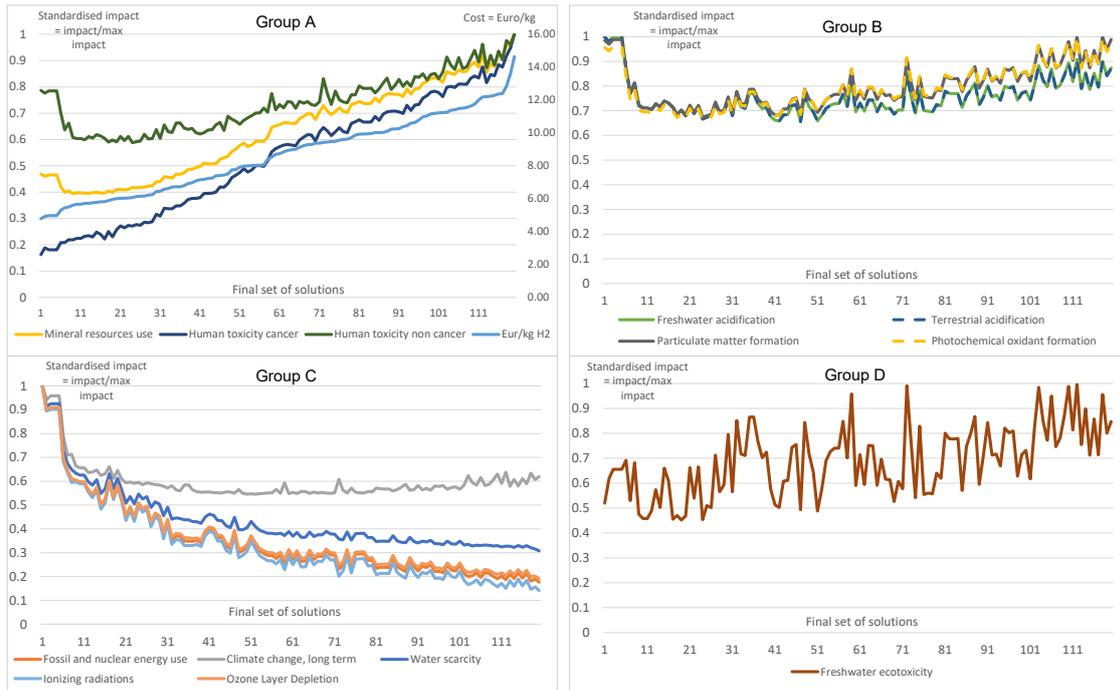


Figure 5.17: The correlations between 4 indicators retained and the other 10 discarded in the final set of 120 solutions

Table 5.7: Correlation coefficients for retained indicators with the discarded ones (except fresh-water ecotoxicity). Full results in table A.13.

Correlation with climate change potential		Correlation with cost		Correlation with Terrestrial acidification
Fossil energy and nuclear use	0.80	Human toxicity cancer	1.00	Freshwater acidification
Water scarcity	0.85	Mineral resource depletion	0.98	Particulate matter formation
Ionising radiation	0.79	Human toxicity non-cancer	0.88	Photochemical oxidant formation
Ozone layer depletion	0.81			

In order to gain insights about the correlations between indicators in figure 5.17, following points are outlined:

Group A (Top left). Overall this group seems to be sensitive to the increase in the sizes of PV and wind. The cost of renewables is higher than the grid for this case study. Similarly for other indicators, impact per kg hydrogen produced is relatively high for renewables.

Group B (Top right). In this group, the relative impact per kg hydrogen produced for all components is the same. See figure 5.6. As a result, for a particular system configuration, each component contributes similarly to the final impact.

Group C (Bottom left). The indicators in this group are highly influential on grid electricity use.

As the grid electricity use decreases from left to right, a decrease in the values is seen. However after solution no. 52, the climate change impact remains constant and slightly increases at the end. This is when further minimisation of grid use oversized solar PV and wind turbines. As a result, much of the electricity is wasted. It leads to decrease in 4 indicators, but all other indicators including cost increase.

Group D (Bottom right). Only freshwater ecotoxicity is included in this group. The comparative impact depends almost entirely on the presence of solar PV. This can also be seen due to relatively high freshwater ecotoxicity per kg hydrogen production as compared to other components in the system. This comes from aluminium and copper emissions in groundwater during mining and treatment of these metals.

Discussion

The proposed approach adheres to the preference of using midpoint indicators for visualization and eventual decision-making. Correlations between the indicators are leveraged for this purpose. A limited set of indicators are provided with correlation information with other indicators in the background. Thus, an overview of the holistic environmental can be taken into account in communication or decision-making.

It is noted that there is a degree of subjectivity in step 3 where preferred indicators and solutions can be added to visualization. However, these choices are often project specific that depend on priorities of decision-makers or collaborators. Additional indicators such as cost/CO₂-eq avoided or assigning preference based weights to indicators such as in multi-criteria decision making can always be added on a project-by-project basis as required.

5.4.3 Comparison with the status-quo

A comparison between the pareto solutions above for 7 objective search (orange points) is made with 2 objective search (blue points) in figure 5.18. As expected, it can be seen that not all solutions that are optimal in the tech-eco-LCA objectives are pareto optimal in the 2 dimensional tech-eco plane. As a result, in the status-quo when search is made only with 2 objectives, when the system is sized with respect to tech-eco objectives, it is not possible to search for solutions with higher environmental efficiency. For example, all the blue points have climate change indicator to approx 4.6 kg CO₂-eq/kg H₂. Max sizing of wind and PV is 0 and 50kW respectively. If for example, climate change impact is desired to be decreased, arbitrary sizing of wind or PV is to be done manually. In this case, it is highly likely that non-optimal solutions will be obtained. That is, it cannot be ensured whether the performance of the system can be improved further. Whereas for pareto-optimal solutions, an objective value can only be improved if at least one of the other objective is made worse.

In this chapter, a systematic approach to include more than 10 LCA indicators in search of solutions, while having a holistic overview of the impact is described. Using the hypervolume indicator and the objective values, pareto optimality of the solutions is highly likely.

Theoretically, all the blue points should also be a part of the orange points in the figure above. Some overlap at the bottom can be seen already. But insufficient coverage of tech-eco points in the overall optimisation is probably since the unmet load was included as a constraint rather than an objective. If required, this can be done easily with additional reference points corresponding to tech-eco minimisation objectives. This will ensure the better coverage of the blue curve.

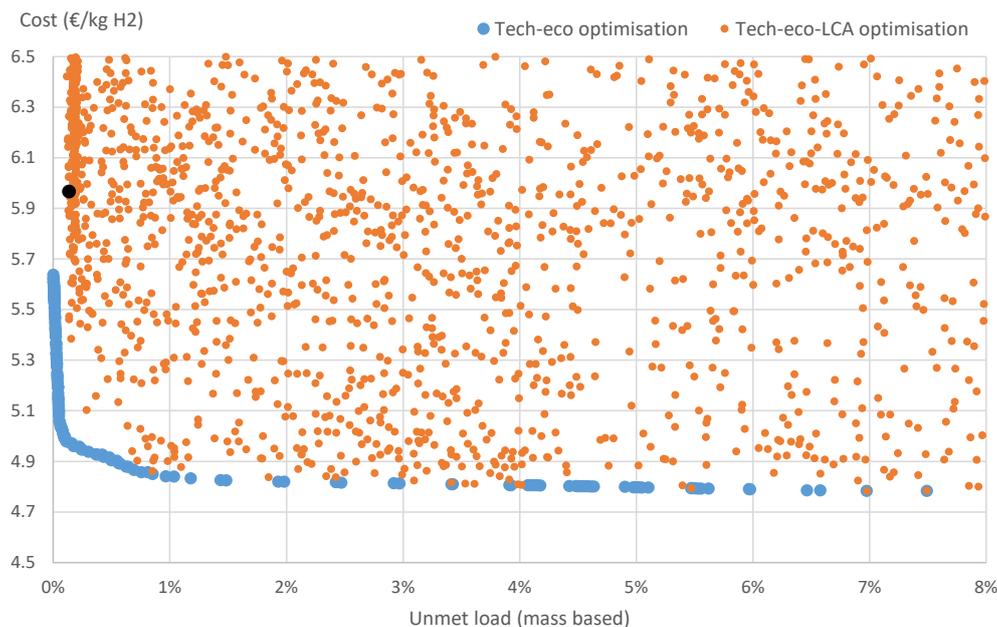


Figure 5.18: Comparison of tech-eco optimisation and with LCA. Black point is solution no. 20 in figure 5.16.

5.5 Conclusions

In this chapter, a systematic approach to take into account overall environmental impact during the search of solutions for hybrid energy systems was presented. Environmental impact is taken into account using LCA midpoint indicators (13), which is arguably the preferred approach of LCA practitioners and personnel at CEA due to its low uncertainty/subjectivity. The presented approach ensures that the LCA practitioner/decision-maker/engineer is at least aware of the impact transfer while searching solutions with higher environmental efficiency. Furthermore, solutions that might be interesting are not skipped by aggregating indicators.

An interesting example regarding its relevance for decision-making is seen when cost per kg H₂ increases from 4.7 € to 14.7 €, climate change potential attains a minimum value of 2.5 kg CO₂-eq/kg H₂ at cost value of around 9 €. Beyond this point, this value remains almost the same even if the cost increases further. This can be seen in figure 5.17 along with relationships between other indicators. Therefore, an approach is provided to facilitate the selection of sensible, technically feasible solutions while considering a holistic environmental impact.

For the multi/many-objective optimisation, one of state-of-the-art genetic algorithms NSGA-III was used. To our knowledge, this is not yet done in the LCA literature. The main advantages being its performance even with upto 15 objectives and preferential search of the pareto front. Consequently, post-processing of pareto front solutions might not be necessary.

The efficiency of NSGA-III was further tested with reduced no. of objectives. It was found that indicator reduction upto a certain point, improves the efficiency of NSGA-III. The main reason being better quality of solutions in terms of both evaluation time and diversity could be obtained. Furthermore, it was found that reduced indicators (except for when objectives were reduced to 2) minimised all objective values in less number of evaluations. Such an investigation of NSGA-III with reduced objectives, to our knowledge is not yet done in the literature.

The limitations of NSGA-III when dealing with high number of objectives were also highlighted. It struggled to obtain a uniform distribution of solutions with 14 objectives. Thus, for utilisation in LCA, where redundant indicators could be easily found, NSGA-III combined with

objective reduction seems promising.

When the number of indicators were reduced to less than 7, a trade-off between time and quality of solutions was observed for NSGA-III. The 14 indicators were reduced to 2 when the framework by [Deb2005, Saxena2012] was followed. However, with the 2 indicators, the performance of the search was found to be unacceptable. LCA being a problem with high number of redundant indicators, most of the variance could be explained by a small set. Nonetheless, leaving out a small amount of variance might considerably worsen the performance of the search. It was found that increasing the CUT value to around 99.99% and not making additional reduction using correlation matrix, maintained the pareto front of the original problem.

An important observation regarding the PCA reduction procedure is that sample set only after 5 generations was selected from a practical viewpoint. This value is much smaller as compared to the literature where often 200 generations are used as a sample set [Saxena2012, Yuan2017]. From the results obtained, it seems to be adequate as a sample set. Perhaps, this case study is an exception or maybe a sample set far from ideal can still be used for objective reduction. Correlation matrices at the end of 5 generations and for the finally visualised 120 solutions can be seen in annex tables A.12, A.13. More research is needed in this direction.

Another observation is that when relying on correlations either for indicator reduction or decision-making, caution is advised since the relationships may change at the extreme points. It was observed even with a correlation coefficient = 1 between two indicators, the relationship deteriorates at their respective minima. This might have an affect during the search of solutions or during the final decision-making. This was also underlined as a drawback of correlation based objective reduction techniques in the investigations of [Yuan2017]. Nonetheless, their advantages far outweigh their disadvantages as compared to other reduction techniques. Higher robustness with non-ideal sample sets and less time required were identified by them (ibid.). One more advantage in terms of correlation based visualization was identified in this chapter. This is not possible with objective reduction techniques such as ones relying on the dominance structure.

Indicator reduction using visual survey was again successfully applied. Correlations between the indicators were identified without even initialising an optimisation problem. As far as understood, this is not possible with other reduction techniques. The indicators classified in the sub-groups had high correlation coefficient between them. It was seen that this could save a significant amount of time required to produce the initial sample set. Furthermore, reliance on the quality of this set can also be eliminated. If the principle of this technique could be automated into an algorithm, its efficiency can be further improved.

General conclusions

At the beginning of this thesis, two main impediments towards including environmental information in the assessments of hybrid energy systems were identified. These two topics were then confronted with the literature to formulate three research proposals that were addressed in this thesis. Using them, a general approach to include environmental information in the assessments of hybrid energy systems is developed as illustrated in figure 5.19. This approach can be divided into 2 parts: parameterized LCA model creation for integrated tech-eco-LCA modelling and inclusion of LCA indicators in multi/many objective optimisation. Each of the two parts are described and their application is justified below.

Integrated tech-eco-LCA modelling (Steps 1-3) Once the project goals are clearly defined, first step is information collection regarding the energy simulations and definition of the LCA goal & scope. The engineers in the laboratory are already well versed with the former since it is their primary activity during the investigations currently carried out in the laboratory. This type of information is usually collected from project stakeholders, databases, literature or technical reports.

For a hybrid energy system, once the LCA goal and scope is defined, the relevant components and process flows under the LCA boundary are identified. Variabilities, uncertainties and data gaps are identified and dealt with, by using initial estimates. These LCA models are coupled with outputs from energy simulations. Parameterised LCA models are thus developed that can take into account these alterations.

Next, the model is improved using global sensitivity analysis. Derivative based global sensitivity measures were used in this thesis since it was found to offer a good trade-off between accuracy in the estimation of sensitivity indices and computational time in this context. Nonetheless, other relevant GSA method could also be used depending on the computational cost and number of uncertainties in the model. Influential parameters and foreground processes are then identified using the obtained sensitivity and importance indices respectively. The model is then improved until the uncertainties are below an acceptable value.

These steps to be followed were presented in detail in chapter 3 and demonstrated with a case study on hydrogen production from PEM water electrolysis. The resulting integrated model is such that it can take in as inputs any parameter and return the corresponding technical, economic and LCA indicators.

Discussion Therefore using this step, the LCA indicators could be calculated with respect to changes in energy simulations such as size, process flows or geography. Thus a link between changes in design (system parameters, configuration) and LCA indicators was established which did not exist before.

Thanks to this inclusion, insights on the environmental performance of the hybrid energy systems can be obtained. These insights could then be used to formulate recommendations in R&D or in planning new facilities in order to improve environmental efficiency. Sometimes the recommendations from an environmental point of view could be opposite from that of an economic perspective. Thus excluding environmental information from assessments could mean that opportunities are missed in increasing environmental efficiency of energy systems.

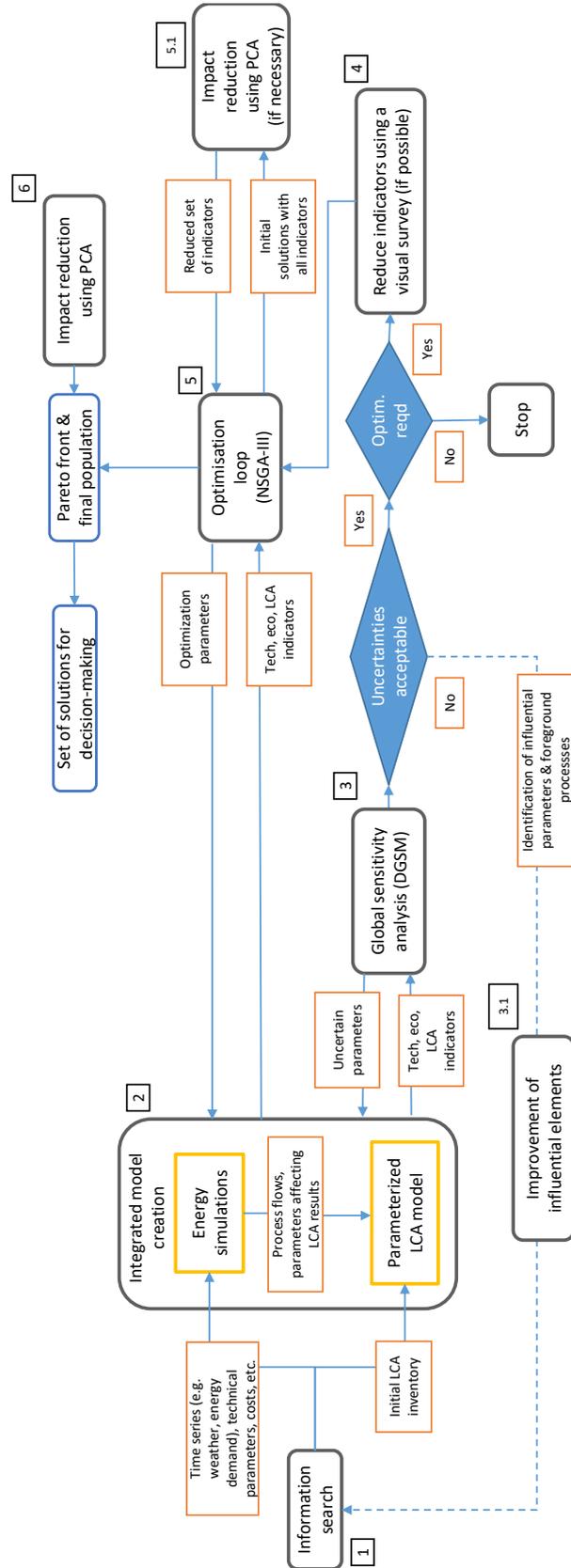


Figure 5.19: Illustration of the proposed approach to include LCA indicators in the assessments of hybrid energy systems

It was found that electricity mix, efficiency and catalyst content in the stack were the most important in the considered environmental indicators. Quantified recommendations regarding these aspects could consequently be provided to decision-makers. Using this information, for example, investment in R&D could be guided for improving the efficiency or specifically reducing the catalyst content in the stack.

Sensitivity and importance indices were also used to guide data collection. Model assembly structure is recommended such that calculation of these coefficients are possible. It allows identifying the parameters and foreground processes that are the most important for improving reliability and robustness of the LCA models.

The proposed framework provides a systematic methodology for building parameterised LCA models. It was developed in tandem with an energy engineer, who was a beginner in LCA. Thus facilitating its use in cases where non-experts in LCA want to add environmental models to hybrid energy simulations. Guidance is provided on various issues that could arise in modelling emerging energy technologies in the future such as: scaling, geographical variability, unavailable data and future evolution.

Optimisation including LCA indicators (Steps 4-6) Then, if optimisation is required, these steps are carried out. An important point when optimisation with LCA is desired is to pre-estimate the difference in the LCA indicators. For example, if the expected difference for a certain indicator between the maximum and minimum values is only 30%, it will be difficult to distinguish the solutions from an LCA point of view. Thus, including LCA indicators in the optimisation might not make sense for this case.

To go ahead with the optimisation, the parameterized tech-eco-LCA model is connected to a new generation evolutionary algorithm, NSGA-III. The reference points are provided as per the desired solutions. They are necessary for functioning of the algorithm. The points 0-1 correspond to the maximum and minimum values of objectives respectively.

Furthermore, it was observed that reduction in indicators improved the quality of solutions and decreased the computational time with NSGA-III. First if possible, reduction of number of indicators by finding correlations between them could be made using a visual survey. Since it will eliminate the requirement of a sample set, saving computational time.

Since the application of this technique is limited, a reduction using principal component analysis could be also be done. Here the initial problem can be run for 5-10 generations and the resulting pareto front can be used as a sample set for reduction. It is accepted that for an ideal sample set, much more generations are required (200-2000 generations in [Saxena2012, Yuan2017]). Then however, objective reduction techniques do not make sense in this case. The number of generations to obtain the sample set could be increased if the computations are not expensive.

As we saw, the results after 5 generations for case study in chapter 5 were promising. Additionally, correlation based objective reduction techniques like PCA, are better equipped to deal with non-ideal sample sets than other types [Yuan2017]. On the other hand, if solutions are selected at 0 or 1st generation, there is a much higher risk that correlations in them do not represent as the ones in the pareto front. This might result in unreliable retention of indicators. A difference between correlations at 5th generation and final set of solutions can be seen in tables A.12 and A.13.

Regarding the PCA reduction procedure outlined by [Deb2005, Saxena2012], following recommendations are made according to the results obtained in chapter 5. If the shape of the original pareto problem is to be maintained, it is recommended that: CUT value should be maintained close to 99.99% to retain principal components and the last step of further reduction using correlation matrix should be skipped. Further reduction might result in a trade-off between solution quality and computational time.

The number of generations for optimisation can be determined by stability in the hypervolume indicator and minimum/average indicator values in the pareto front. The final population and

the pareto optimal solutions can then be used for visualisation of solutions for decision-making. Since the final population is representative of the preferred solutions, they are a valid starting point. If they are not optimal, they are replaced by similar individuals in the obtained pareto front using euclidean distance. At this point, PCA is recommended to be conducted again but a lower CUT value $\approx 95\%$ to select indicators could be afforded. Here, PCA is used to ease the visualization, thus the exact CUT value can be decided the users themselves to retain indicators. These indicators are then used to represent the final solution set. The correlation information of the discarded indicators is provided with the retained indicators.

This is an iterative step where other solutions from the pareto could be added using euclidean distance or using additional criteria such fixed cost. Thus it is possible to make a decision based of few indicators while having a holistic overview of the environmental impacts. The proposed approach thus leverages the operating principle of NSGA-III and the correlation based objective reduction techniques.

Discussion This step enables the search of system configurations with higher environmental efficiency. It was shown that this is not possible if optimisation is made only with techno-economic indicators. For instance for the techno-economic optimisation, all solutions had the almost same climate change potential ≈ 4.6 kg CO₂-eq/kg H₂. Including LCA indicators enabled search of solutions upto 2.5 kg CO₂-eq/kg H₂ along with trade-offs with technical and other LCA indicators. In chapter 5, this approach was applied to a dimensioning problem of hydrogen production from PEM water electrolysis. A pareto front with 1 economic and 13 LCA midpoint indicators was obtained, while technical performance was added as a constraint. This offers a choice of solutions with higher environmental efficiency to the users and decision-makers. Trade-offs between these techno-economic and environmental indicators could be found. Additionally, even the trade-offs between the LCA indicators themselves could be obtained. Thus, the impact transfer can be kept in check.

For multi-objective optimisation, evolutionary algorithm NSGA-III coupled with objective reduction techniques is proposed. To the best of our knowledge, it has not been done yet in the literature. It was shown that performance of NSGA-III could be improved using objective reduction techniques. Higher quality of solutions could be obtained in less time. For example, computational time was reduced to half when 14 indicators were reduced to 7 using correlations between them while simultaneously obtaining higher hypervolume indicator value.

Furthermore, by changing settings in the PCA reduction procedure and using the new visual technique for indicator reduction, a trade-off between solution quality and computational time was revealed when indicators were reduced less than 7. Regarding the PCA reduction procedure, when the framework proposed by the authors [Deb2005, Saxena2012] was followed, it led to retaining only 2 objectives that gave the worst pareto solution quality in terms of convergence. In order to maintain the structure of original pareto problem, high threshold value $\approx 99.9\%$ is recommended while additional reduction in the last step using a correlation matrix is not recommended.

A new indicator reduction using a visual technique was found during the implementation of the above algorithm. It showed promising results in the identification of correlations between LCA indicators without requiring samples. It can be applied to simple cases but its efficacy could be improved in the future.

Perspectives

Sample sets for objective reduction techniques. Initialization of a problem in all objective space is required to generate samples. Objective reduction techniques rely on these samples for their selections. Some examples of this duration from the literature are: 20 minutes in [Brockhoff2007], 15.6 hours/400 generations in [Perez Gallardo2013], 200 - 2000 generations in [Yuan2017, Saxena2012]. This is done so that the correlations or dominance structure in the sample set reflect that of the pareto front for reliable reduction of indicators. The problem is, when a functional evaluation is expensive, like in this thesis, the objective reduction techniques will increase the computational time instead of decreasing it. In this thesis, samples were selected for PCA objective reduction after 5 generations from a practical point of view, which gave promising results. Limited research was found on the duration of this initialization step, as was also concluded by [Yuan2017]. Thus question can be asked, what is the sampling duration for objective reduction in a reasonable time?

Visual survey as an algorithm. The visual technique of objective reduction correctly identified the correlations between the LCA indicators without requiring a sampling set. This was demonstrated for two cases. Nevertheless, it is only a principle at this stage that could be implemented into an algorithm for objective reduction. The main advantage is that obtaining initial samples as mentioned in the above paragraph, is not necessary. By implementation and testing, its application could be refined. The elements in this technique that could be refined include: range for classification of indicators in sub-groups (step 4 in figure 4.3) and dealing with complex/non-linear cases.

Treatment of uncertainties. In the framework for developing LCA models, uncertainties or variabilities due to geography, data gaps, size and different process/technologies were taken into account. Other uncertainties, although not in the scope of thesis, could play an important role in the assessments. Improving them would ensure derivation of more robust conclusions. The framework could be extended to include these uncertainties. Another interesting avenue would be to include uncertainties in the optimisation process as well. A recent method to deal with techno-economic uncertainties in the optimisation of hybrid energy systems could be extended to also include LCA indicators [Nadal2020].

Online objective reduction with NSGA-III. Objective reduction techniques integrated into an evolutionary algorithm itself are possible. Also called, *online* objective reduction, they have been implemented with NSGA-II [Cheung2014]. An online integration of PCA with NSGA-III could offer interesting results.

Repeatability. Genetic algorithms are stochastic methods that start with a randomly initialized population. Furthermore, in this case the number of optimal individuals are quite high since pareto front is a curve in 14 dimensions. Out of them a finite number of solutions are found and consequently selected for decision making. Thus it is possible that different solutions are obtained at the end of each optimisation run. It will be interesting to assess the differences between the

obtained solutions at the end of each optimisation run. In terms of system configurations as well as indicator values. Then consequently exploring if there are strategies to increase repeatability of optimisation results.

Application to different case studies. From a more practical point of view, application of the optimisation approach to more case studies and energy systems could offer promising results. An example could be inclusion of LCA indicators in the optimisation of an electricity grid or planning investment of energy technologies for the future. Instances where LCA indicators are not considered as relevant criteria for design and planning. It could lead to new solutions that were not visible before. Development of a user interface of the concepts introduced in this thesis will greatly facilitate its application to different case studies by making it easier to use.

Abbreviations · Publications

Abbreviations

ANN	Artificial neural networks
BOP	Balance of plant
CEA	French commission of atomic and alternative energies
CUT	Threshold value of cumulative variance of PCs
DEAP	Distributed evolutionary algorithms in python
DGSM	Derivative based global sensitivity measure
EF	Environmental footprint impact assessment method
EOL	End of life
EU	European union
GSA	Global sensitivity analysis
ILCD	International Reference Life Cycle Data System
LCA	Life cycle assessment
LCI	Life cycle inventory analysis
LSA	Local sensitivity analysis
LSED	Laboratory of energy systems and demonstrations, CEA-Liten
MILP	Multi integer linear programming
NREL	National Renewable Energy Laboratory
PC	Principal component
PCA	Principal component analysis
PEM	Polymer electrolyte membrane
PV	Photovoltaics
R&D	Research and development
SALib	Sensitivity analysis library in python
TRL	Technology readiness level

Conference publications

Sharma, H., Mandil, G., Zwolinski, P., Cor, E., Mugnier, H., & Monnier, E. (2020). Integration of life cycle assessment with energy simulation software for polymer exchange membrane (PEM) electrolysis. *Procedia CIRP*, 90, 176-181. 27th life cycle engineering conference, Grenoble, France. May 2020. DOI: <https://doi.org/10.1016/j.procir.2020.02.139>

Sharma, H., Monnier, É., Mandil, G., Zwolinski, P., & Colasson, S. (2019). Comparison of environmental assessment methodology in hybrid energy system simulation software. *Procedia CIRP*, 80, 221-227. 26th life cycle engineering conference, Purdue, USA. May 2019. DOI: <https://doi.org/10.1016/j.procir.2019.01.007>

Bibliography

- [ADEME2014] ADEME. *ADEME ENERGY TRANSITION SCENARIOS 2030 - 2050*. Rapport technique, November 2014.
- [ADEME2018] ADEME. *Déchets chiffres-clés - L'essentiel 2018*, 2018.
- [Ahmadi2015] Aras Ahmadi et Ligia Tiruta-Barna. *A Process Modelling-Life Cycle Assessment-MultiObjective Optimization tool for the eco-design of conventional treatment processes of potable water*. Journal of Cleaner Production, vol. 100, pages 116–125, 2015.
- [Antipova2013] Ekaterina Antipova, Dieter Boer, Luisa F Cabeza, Gonzalo Guillén-Gosálbez et Laureano Jiménez. *Uncovering relationships between environmental metrics in the multi-objective optimization of energy systems: A case study of a thermal solar Rankine reverse osmosis desalination plant*. Energy, vol. 51, pages 50–60, 2013.
- [Antipova2014] Ekaterina Antipova, Dieter Boer, Gonzalo Guillén-Gosálbez, Luisa F Cabeza et Laureano Jiménez. *Multi-objective optimization coupled with life cycle assessment for retrofitting buildings*. Energy and buildings, vol. 82, pages 92–99, 2014.
- [Antipova2015] Ekaterina Antipova, Carlos Pozo, Gonzalo Guillén-Gosálbez, Dieter Boer, Luisa F Cabeza et Laureano Jiménez. *On the use of filters to facilitate the post-optimal analysis of the Pareto solutions in multi-objective optimization*. Computers & Chemical Engineering, vol. 74, pages 48–58, 2015.
- [Arvidsson2018] Rickard Arvidsson, Anne-Marie Tillman, Björn A. Sandén, Matty Janssen, Anders Nordelöf, Duncan Kushnir et Sverker Molander. *Environmental Assessment of Emerging Technologies: Recommendations for Prospective LCA*. Journal of Industrial Ecology, vol. 22, no. 6, pages 1286–1294, December 2018.
- [Azapagic1999] Adisa Azapagic et R Clift. *Life cycle assessment and multiobjective optimisation*. Journal of cleaner production, vol. 7, no. 2, pages 135–143, 1999.
- [Bare2000] Jane C Bare, Patrick Hofstetter, David W Pennington et Helias A Udo De Haes. *Midpoints versus endpoints: the sacrifices and benefits*. The International Journal of Life Cycle Assessment, vol. 5, no. 6, page 319, 2000.
- [Bare2006] Jane C Bare et Thomas P Gloria. *Critical analysis of the mathematical relationships and comprehensiveness of life cycle impact assessment approaches*. Environmental science & technology, vol. 40, no. 4, pages 1104–1113, 2006.

- [Barei2019] Kay Barei, Cristina de la Rua, Maximilian Mckl et Thomas Hamacher. *Life cycle assessment of hydrogen from proton exchange membrane water electrolysis in future energy systems*. Applied Energy, vol. 237, pages 862–872, March 2019.
- [Beccali2016] Marco Beccali, Maurizio Cellura, Sonia Longo et Francesco Guarino. *Solar heating and cooling systems versus conventional systems assisted by photovoltaic: Application of a simplified LCA tool*. Solar Energy Materials and Solar Cells, vol. 156, pages 92–100, 2016.
- [Bechikh2016] Slim Bechikh, Rituparna Datta et Abhishek Gupta. Recent advances in evolutionary multi-objective optimization, volume 20. Springer, 2016.
- [Bergerson2020] Joule A. Bergerson, Adam Brandt, Joe Cresko, Michael Carbajales-Dale, Heather L. MacLean, H. Scott Matthews, Sean McCoy, Marcelle McManus, Shelie A. Miller, William R. Morrow, I. Daniel Posen, Thomas Seager, Timothy Skone et Sylvia Sleep. *Life cycle assessment of emerging technologies: Evaluation techniques at different stages of market and technical maturity*. Journal of Industrial Ecology, vol. 24, no. 1, pages 11–25, February 2020.
- [Bergesen2016] Joseph D. Bergesen et Sangwon Suh. *A framework for technological learning in the supply chain: A case study on CdTe photovoltaics*. Applied Energy, vol. 169, pages 721–728, May 2016.
- [Bertuccioli2014] Luca Bertuccioli, Alvin Chan, David Hart, Franz Lehner, Ben Madden et Eleanor Standen. *Development of water electrolysis in the European Union*. Fuel cells and hydrogen joint undertaking, vol. 83, 2014.
- [Bhandari2014] Ramchandra Bhandari, Clemens A. Trudewind et Petra Zapp. *Life cycle assessment of hydrogen production via electrolysis – a review*. Journal of Cleaner Production, vol. 85, pages 151–163, December 2014.
- [Bjrklund2002] Anna E Bjrklund. *Survey of approaches to improve reliability in LCA*. The International Journal of Life Cycle Assessment, vol. 7, no. 2, page 64, 2002.
- [Bramer2007] Max Bramer. Principles of data mining, volume 180. Springer, 2007.
- [Brockhoff2006] Dimo Brockhoff et Eckart Zitzler. *Are all objectives necessary? On dimensionality reduction in evolutionary multiobjective optimization*. In Parallel Problem Solving from Nature-PPSN IX, pages 533–542. Springer, 2006.
- [Brockhoff2007] Dimo Brockhoff et Eckart Zitzler. *Improving hypervolume-based multiobjective evolutionary algorithms by using objective reduction methods*. In 2007 IEEE congress on evolutionary computation, pages 2086–2093. IEEE, 2007.

-
- [Brondizio2019] ES Brondizio, J Settele, S Díaz, HT Ngoet *al.* *Global assessment report on biodiversity and ecosystem services*. The Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services, Bonn, Germany, 2019.
- [Brunet2012] Robert Brunet, Gonzalo Guillén-Gosálbez et Laureano Jiménez. *Cleaner design of single-product biotechnological facilities through the integration of process simulation, multiobjective optimization, life cycle assessment, and principal component analysis*. *Industrial & Engineering Chemistry Research*, vol. 51, no. 1, pages 410–424, 2012.
- [Bulle2019] Cécile Bulle, Manuele Margni, Laure Patouillard, Anne-Marie Boulay, Guillaume Bourgault, Vincent De Bruille, Viêt Cao, Michael Hauschild, Andrew Henderson, Sebastien Humbert, Sormeh Kashef-Haghighi, Anna Kounina, Alexis Laurent, Annie Levasseur, Gladys Liard, Ralph K. Rosenbaum, Pierre-Olivier Roy, Shanna Shaked, Peter Fantke et Olivier Jolliet. *IMPACT World+: a globally regionalized life cycle impact assessment method*. *The International Journal of Life Cycle Assessment*, vol. 24, no. 9, pages 1653–1674, September 2019.
- [Buttler2018] Alexander Buttler et Hartmut Spliethoff. *Current status of water electrolysis for energy storage, grid balancing and sector coupling via power-to-gas and power-to-liquids: A review*. *Renewable and Sustainable Energy Reviews*, vol. 82, pages 2440–2454, February 2018.
- [Buyle2019] Buyle, Audenaert, Billen, Boonen et Van Passel. *The Future of Ex-Ante LCA? Lessons Learned and Practical Recommendations*. *Sustainability*, vol. 11, no. 19, page 5456, October 2019.
- [Caduff2011] Marloes Caduff, Mark A. J. Huijbregts, Hans-Joerg Althaus et A. Jan Hendriks. *Power-Law Relationships for Estimating Mass, Fuel Consumption and Costs of Energy Conversion Equipments*. *Environmental Science & Technology*, vol. 45, no. 2, pages 751–754, January 2011.
- [Caduff2012] Marloes Caduff, Mark AJ Huijbregts, Hans-Joerg Althaus, Annette Koehler et Stefanie Hellweg. *Wind power electricity: the bigger the turbine, the greener the electricity?* *Environmental science & technology*, vol. 46, no. 9, pages 4725–4733, 2012.
- [Caduff2014] Marloes Caduff, Mark A.J. Huijbregts, Annette Koehler, Hans-Jörg Althaus et Stefanie Hellweg. *Scaling Relationships in Life Cycle Assessment: The Case of Heat Production from Biomass and Heat Pumps*. *Journal of Industrial Ecology*, vol. 18, no. 3, pages 393–406, May 2014.
- [Carlson2005] E. J. Carlson, P. Kopf, J. Sinha, S. Sriramulu et Y. Yang. *Cost Analysis of PEM Fuel Cell Systems for Transportation: September 30, 2005*. Rapport technique NREL/SR-560-39104, 862302, December 2005.
-

- [Carmo2013] Marcelo Carmo, David L. Fritz, Jürgen Mergel et Detlef Stolten. *A comprehensive review on PEM water electrolysis*. International Journal of Hydrogen Energy, vol. 38, no. 12, pages 4901–4934, April 2013.
- [Ceballos2015] Gerardo Ceballos, Paul R Ehrlich, Anthony D Barnosky, Andrés García, Robert M Pringle et Todd M Palmer. *Accelerated modern human–induced species losses: Entering the sixth mass extinction*. Science advances, vol. 1, no. 5, page e1400253, 2015.
- [Chauvel2003] Alain Chauvel, Gilles Fournier et Claude Raimbault. *Manual of process economic evaluation*. Editions Technip, 2003.
- [Cheung2014] Yiu-ming Cheung et Fangqing Gu. *Online objective reduction for many-objective optimization problems*. In 2014 IEEE congress on evolutionary computation (CEC), pages 1165–1171. IEEE, 2014.
- [Ciroth2014] Andreas Ciroth, Sarah Winter et GreenDelta Berlin. *openLCA 1.4 overview and first steps*. GreenDelta, Berlin, 2014.
- [Connolly2010] D. Connolly, H. Lund, B.V. Mathiesen et M. Leahy. *A review of computer tools for analysing the integration of renewable energy into various energy systems*. vol. 87, no. 4, pages 1059–1082, 2010.
- [Cucek2013] Lidija Cucek, Jiri Jaromir Klemes, Petar Sabev Varbanov et Zdravko Kravanja. *Dealing with high-dimensionality of criteria in multiobjective optimization of biomass energy supply network*. Industrial & Engineering Chemistry Research, vol. 52, no. 22, pages 7223–7239, 2013.
- [Čuček2014] Lidija Čuček, Jiří Jaromír Klemeš et Zdravko Kravanja. *Objective dimensionality reduction method within multi-objective optimisation considering total footprints*. Journal of cleaner production, vol. 71, pages 75–86, 2014.
- [Das1998] Indraneel Das et John E Dennis. *Normal-boundary intersection: A new method for generating the Pareto surface in nonlinear multicriteria optimization problems*. SIAM journal on optimization, vol. 8, no. 3, pages 631–657, 1998.
- [Deb2005] Kalyanmoy Deb et Dhish Kumar Saxena. *On finding pareto-optimal solutions through dimensionality reduction for certain large-dimensional multi-objective optimization problems*. Kangal report, vol. 2005011, 2005.
- [Deb2013] Kalyanmoy Deb et Himanshu Jain. *An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: solving problems with box constraints*. IEEE transactions on evolutionary computation, vol. 18, no. 4, pages 577–601, 2013.
- [Deng2017] Gary Deng et Peter Newton. *Assessing the impact of solar PV on domestic electricity consumption: Exploring the prospect of rebound effects*. Energy Policy, vol. 110, pages 313–324, 2017.

-
- [Di Lullo2020] Giovanni Di Lullo, Eskinder Gemechu, Abayomi Olufemi Oni et Amit Kumar. *Extending sensitivity analysis using regression to effectively disseminate life cycle assessment results*. The International Journal of Life Cycle Assessment, vol. 25, no. 2, pages 222–239, 2020.
- [Dietz2006] Adrian Dietz, Catherine Azzaro-Pantel, Luc Pibouleau et Serge Domenech. *Multiobjective optimization for multiproduct batch plant design under economic and environmental considerations*. Computers & Chemical Engineering, vol. 30, no. 4, pages 599–613, 2006.
- [Dong2014] Ya Hong Dong et S Thomas Ng. *Comparing the midpoint and endpoint approaches based on ReCiPe—a study of commercial buildings in Hong Kong*. The International Journal of Life Cycle Assessment, vol. 19, no. 7, pages 1409–1423, 2014.
- [Douglas-Smith2020] Dominique Douglas-Smith, Takuya Iwanaga, Barry FW Croke et Anthony J Jakeman. *Certain trends in uncertainty and sensitivity analysis: An overview of software tools and techniques*. Environmental Modelling & Software, vol. 124, page 104588, 2020.
- [Duclos2017] Lucien Duclos, Maria Lupsea, Guillaume Mandil, Lenka Svecova, Pierre-Xavier Thivel et Valérie Laforest. *Environmental assessment of proton exchange membrane fuel cell platinum catalyst recycling*. Journal of Cleaner Production, vol. 142, pages 2618–2628, January 2017.
- [European Commission2010] European Commission et Joint Research Centre. *ILCD handbook: general guide for life cycle assessment : detailed guidance*. Publications Office of the European Union, Luxembourg, 2010. OCLC: 711747737.
- [EuropeanCommission2011] EuropeanCommission. *International Reference Life Cycle Data System (ILCD) Handbook- Recommendations for Life Cycle Impact Assessment in the European context. First edition November 2011. EUR 24571 EN. Luxemburg. Publications Office of the European Union; 2011*. European Commission-Joint Research Centre, 2011.
- [Evangelisti2017] Sara Evangelisti, Carla Tagliaferri, Dan JL Brett et Paola Lettieri. *Life cycle assessment of a polymer electrolyte membrane fuel cell system for passenger vehicles*. Journal of cleaner production, vol. 142, pages 4339–4355, 2017.
- [Fazio2018] S Fazio, F Biganzioli, V De Laurentiis, L Zampori, S Sala et E Diaconu. *Supporting information to the characterisation factors of recommended EF Life Cycle Impact Assessment methods, version 2, from ILCD to EF 3.0*. Ispra. doi, vol. 10, page 002447, 2018.
- [Fonseca2006] Carlos M Fonseca, Luís Paquete et Manuel López-Ibáñez. *An improved dimension-sweep algorithm for the hypervolume indicator*. In 2006 IEEE international conference on evolutionary computation, pages 1157–1163. IEEE, 2006.
-

- [Fortin2012] Félix-Antoine Fortin, François-Michel De Rainville, Marc-André Gardner Gardner, Marc Parizeau et Christian Gagné. *DEAP: Evolutionary algorithms made easy*. The Journal of Machine Learning Research, vol. 13, no. 1, pages 2171–2175, 2012.
- [Gantner2018] Johannes Gantner, Katrin Lenz, Rafael Horn, Petra Von Both et Sebastian Ebertshäuser. *Ökobau. dat 3.0–Quo Vadis? Buildings*, vol. 8, no. 9, page 129, 2018.
- [Gazbour2016] Nouha Gazbour, Guillaume Razongles, Christian Schaeffer et Carole Charbuillet. *Photovoltaic power goes green*. In 2016 Electronics Goes Green 2016+(EGG), pages 1–8. IEEE, 2016.
- [Gazbour2018] Nouha Gazbour, Guillaume Razongles, Elise Monnier, Maryline Joanny, Carole Charbuillet, Françoise Burgun et Christian Schaeffer. *A path to reduce variability of the environmental footprint results of photovoltaic systems*. Journal of Cleaner Production, vol. 197, pages 1607–1618, October 2018.
- [Gazbour2019] Nouha Gazbour. *Systemic integration of Eco-design in photovoltaic technologies R&D*. Theses, Université Grenoble Alpes, February 2019. Issue: 2019GREAT018.
- [Gerber2011] Léda Gerber, Martin Gassner et François Maréchal. *Systematic integration of LCA in process systems design: Application to combined fuel and electricity production from lignocellulosic biomass*. Computers & Chemical Engineering, vol. 35, no. 7, pages 1265–1280, July 2011.
- [Gerber2012] Léda Gerber. *Integration of Life Cycle Assessment in the conceptual design of renewable energy conversion systems*. Thèse de Doctorat, ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE, Lausanne, 2012.
- [Gonzalez-Garay2018] Andres Gonzalez-Garay et Gonzalo Guillen-Gosalbez. *SUSCAPE: A framework for the optimal design of SUSTainable ChemicAl ProcEsses incorporating data envelopment analysis*. Chemical Engineering Research and Design, vol. 137, pages 246–264, 2018.
- [Groen2017] Evelyne A. Groen, Eddie A. M. Bokkers, Reinout Heijungs et Imke J. M. de Boer. *Methods for global sensitivity analysis in life cycle assessment*. The International Journal of Life Cycle Assessment, vol. 22, no. 7, pages 1125–1137, July 2017.
- [Guillén-Gosálbez2011] Gonzalo Guillén-Gosálbez. *A novel MILP-based objective reduction method for multi-objective optimization: Application to environmental problems*. Computers & Chemical Engineering, vol. 35, no. 8, pages 1469–1477, 2011.
- [Guinot2013] Benjamin Guinot. *Multi-criteria assessment of storage technologies of renewable energies : Design and realisation of the simulation platform ODYSSEY for the optimisation of sizing and energy control strategies*. Theses, Université de Grenoble, September 2013.

-
- [Guinot2015a] Benjamin Guinot, Bénédicte Champel, Florent Montignac, Elisabeth Lemaire, Didier Vannucci, Sebastien Sailer et Yann Bultel. *Techno-economic study of a PV-hydrogen-battery hybrid system for off-grid power supply: Impact of performances' ageing on optimal system sizing and competitiveness*. International Journal of Hydrogen Energy, vol. 40, no. 1, pages 623–632, 2015.
- [Guinot2015b] Benjamin Guinot, Florent Montignac, Bénédicte Champel et Didier Vannucci. *Profitability of an electrolysis based hydrogen production plant providing grid balancing services*. International Journal of Hydrogen Energy, vol. 40, no. 29, pages 8778–8787, 2015.
- [Gutiérrez2010] Ester Gutiérrez, Sebastián Lozano et Belarmino Adenso-Díaz. *Dimensionality reduction and visualization of the environmental impacts of domestic appliances*. Journal of Industrial Ecology, vol. 14, no. 6, pages 878–889, 2010.
- [Hassim2010] Mimi H. Hassim, Alberto L. Pérez et Markku Hurme. *Estimation of chemical concentration due to fugitive emissions during chemical process design*. Process Safety and Environmental Protection, vol. 88, no. 3, pages 173–184, May 2010.
- [Hauschild2013] Michael Z Hauschild, Mark Goedkoop, Jeroen Guinée, Reinout Heijungs, Mark Huijbregts, Olivier Jolliet, Manuele Margni, An De Schryver, Sebastien Humbert, Alexis Laurent *et al.* *Identifying best existing practice for characterization modeling in life cycle impact assessment*. The International Journal of Life Cycle Assessment, vol. 18, no. 3, pages 683–697, 2013.
- [Hauschild2015] Michael Z Hauschild et Mark AJ Huijbregts. *Introducing life cycle impact assessment*. In Life cycle impact assessment, pages 1–16. Springer, 2015.
- [Heijungs2002] Reinout Heijungs et Sangwon Suh. *The computational structure of life cycle assessment*, volume 11. Springer Science & Business Media, 2002.
- [Heijungs2020] Reinout Heijungs. *Is mainstream LCA linear?* The International Journal of Life Cycle Assessment, vol. 25, no. 10, pages 1872–1882, 2020.
- [Hellweg2014] Stefanie Hellweg et Llorenç Milà i Canals. *Emerging approaches, challenges and opportunities in life cycle assessment*. Science, vol. 344, no. 6188, pages 1109–1113, 2014.
- [Hennen2017] Maike Hennen, Sarah Postels, Philip Voll, Matthias Lampe et André Bardow. *Multi-objective synthesis of energy systems: Efficient identification of design trade-offs*. Computers & Chemical Engineering, vol. 97, pages 283–293, 2017.
- [Herman2017] Jon Herman et Will Usher. *SALib: An open-source Python library for Sensitivity Analysis*. The Journal of Open Source Software, vol. 2, no. 9, page 97, January 2017.
- [Herzog2009] Tim Herzog. *World greenhouse gas emissions in 2005*. World Resources Institute, 2009.
-

- [Hetherington2014] Alexandra C. Hetherington, Aiduan Li Borrión, Owen Glyn Griffiths et Marcelle C. McManus. *Use of LCA as a development tool within early research: challenges and issues across different sectors*. The International Journal of Life Cycle Assessment, vol. 19, no. 1, pages 130–143, January 2014.
- [Hischier2005] Roland Hischier, Stefanie Hellweg, Christian Capello et Alex Primas. *Establishing Life Cycle Inventories of Chemicals Based on Differing Data Availability (9 pp)*. The International Journal of Life Cycle Assessment, vol. 10, no. 1, pages 59–67, January 2005.
- [Hossain2014] M Shahriar Hossain, Manish Marwah, Amip Shah, Layne T Watson et Naren Ramakrishnan. *AutoLCA: a framework for sustainable redesign and assessment of products*. ACM Transactions on Intelligent Systems and Technology (TIST), vol. 5, no. 2, pages 1–21, 2014.
- [Huang2016] C-Y Huang, AH Hu, J Yin et H-C Wang. *Developing a parametric carbon footprinting tool for the semiconductor industry*. International journal of environmental science and technology, vol. 13, no. 1, pages 275–284, 2016.
- [Huijbregts1998] Mark AJ Huijbregts. *Application of uncertainty and variability in LCA*. The International Journal of Life Cycle Assessment, vol. 3, no. 5, page 273, 1998.
- [Huijbregts2016] MAJ Huijbregts, ZJN Steinmann, PMF Elshout, G Stam, F Verones, MDM Vieira, A Hollander, M Zijp et R Van Zelm. *ReCiPe 2016: a harmonized life cycle impact assessment method at midpoint and endpoint level report I: characterization*. 2016.
- [Iguar2017] Laura Iguar et Santi Seguí. *Introduction to Data Science*. In Introduction to Data Science, pages 1–4. Springer, 2017.
- [Iooss2012] B Iooss, AL Popelin, G Blatman, C Ciric, F Gamboa, S Lacaze et M Lamboni. *Some new insights in derivative-based global sensitivity measures*. In Proceedings of the PSAM11 ESREL 2012 Conference, pages 1094–1104, 2012.
- [Iooss2017] Bertrand Iooss et Andrea Saltelli. *Introduction to Sensitivity Analysis*. In Roger Ghanem, David Higdon et Houman Owhadi, editors, Handbook of Uncertainty Quantification, pages 1103–1122. Springer International Publishing, Cham, 2017.
- [Iooss2019] Bertrand Iooss, A Janon, Gilles Pujol, B Broto, K Boumhaout, S Da Veiga, T Delage, J Fruth, L Gilquin, J Guillaume et al. *Sensitivity: global sensitivity analysis of model outputs*. R package version, vol. 1, no. 2, 2019.
- [IPCC2014] IPCC. *Climate change 2014: Synthesis report. Summary for policy-makers*. Contribution of working groups I, II and III to the fifth assessment report of the Intergovernmental panel on climate change, page 151, 2014.

-
- [Ishibuchi2018] Hisao Ishibuchi, Ryo Imada, Yu Setoguchi et Yusuke Nojima. *How to Specify a Reference Point in Hypervolume Calculation for Fair Performance Comparison*. *Evolutionary Computation*, vol. 26, no. 3, pages 411–440, 2018.
- [Jain2013] Himanshu Jain et Kalyanmoy Deb. *An evolutionary many-objective optimization algorithm using reference-point based nondominated sorting approach, part II: handling constraints and extending to an adaptive approach*. *IEEE Transactions on evolutionary computation*, vol. 18, no. 4, pages 602–622, 2013.
- [Jeong2014] Myeon-Gyu Jeong, James R Morrison et Hyo-Won Suh. *Approximate life cycle assessment via case-based reasoning for eco-design*. *IEEE Transactions on Automation Science and Engineering*, vol. 12, no. 2, pages 716–728, 2014.
- [Kägi2016] Thomas Kägi, Fredy Dinkel, Rolf Frischknecht, Sebastien Humbert, Jacob Lindberg, Steven De Mester, Tommie Ponsioen, Serenella Sala et Urs Walter Schenker. *Session “Midpoint, endpoint or single score for decision-making?”—SETAC Europe 25th Annual Meeting, May 5th, 2015*. *The International Journal of Life Cycle Assessment*, vol. 21, no. 1, pages 129–132, 2016.
- [Kalogirou2001] Soteris A Kalogirou. *Artificial neural networks in renewable energy systems applications: a review*. *Renewable and sustainable energy reviews*, vol. 5, no. 4, pages 373–401, 2001.
- [Karka2019] Paraskevi Karka, Stavros Papadokonstantakis et Antonis Kokosis. *Environmental impact assessment of biomass process chains at early design stages using decision trees*. *The International Journal of Life Cycle Assessment*, vol. 24, no. 9, pages 1675–1700, 2019.
- [Khanali2017] Majid Khanali, Hossein Mobli et Homa Hosseinzadeh-Bandbafha. *Modeling of yield and environmental impact categories in tea processing units based on artificial neural networks*. *Environmental Science and Pollution Research*, vol. 24, no. 34, pages 26324–26340, 2017.
- [Kiparissides2009] A. Kiparissides, S. S. Kucherenko, A. Mantalaris et E. N. Pistikopoulos. *Global Sensitivity Analysis Challenges in Biological Systems Modeling*. *Industrial & Engineering Chemistry Research*, vol. 48, no. 15, pages 7168–7180, August 2009.
- [Kleinekorte2019] Johanna Kleinekorte, Leif Kröger, Kai Leonhard et André Bardow. *A neural network-based framework to predict process-specific environmental impacts*. *Computer Aided Chemical Engineering*, vol. 46, pages 1447–1452, 2019.
- [Kopp2017] M. Kopp, D. Coleman, C. Stiller, K. Scheffer, J. Aichinger et B. Scheppat. *Energiepark Mainz: Technical and economic analysis of the worldwide largest Power-to-Gas plant with PEM electrolysis*. *International Journal of Hydrogen Energy*, vol. 42, no. 19, pages 13311–13320, May 2017.
-

- [Kostin2012] A Kostin, G Guillén-Gosálbez, FD Mele et L Jiménez. *Identifying key life cycle assessment metrics in the multiobjective design of bioethanol supply chains using a rigorous mixed-integer linear programming approach*. Industrial & Engineering Chemistry Research, vol. 51, no. 14, pages 5282–5291, 2012.
- [Kucherenko2017] Sergey Kucherenko et Bertrand Iooss. *Derivative-Based Global Sensitivity Measures*. In Roger Ghanem, David Higdon et Houman Owhadi, éditeurs, Handbook of Uncertainty Quantification, pages 1241–1263. Springer International Publishing, Cham, 2017.
- [Lacirignola2017] Martino Lacirignola, Philippe Blanc, Robin Girard, Paula Perez-Lopez et Isabelle Blanc. *LCA of emerging technologies: addressing high uncertainty on inputs' variability when performing global sensitivity analysis*. Science of the Total Environment, vol. 578, pages 268–280, 2017.
- [Lamboni2013] M. Lamboni, B. Iooss, A.-L. Popelin et F. Gamboa. *Derivative-based global sensitivity measures: General links with Sobol' indices and numerical tests*. Mathematics and Computers in Simulation, vol. 87, pages 45–54, January 2013.
- [Laurent2012] Alexis Laurent, Stig I Olsen et Michael Z Hauschild. *Limitations of carbon footprint as indicator of environmental sustainability*. Environmental science & technology, vol. 46, no. 7, pages 4100–4108, 2012.
- [Lenzen2006] Manfred Lenzen. *Uncertainty in impact and externality assessments-implications for decision-making (13 pp)*. The International journal of life cycle assessment, vol. 11, no. 3, pages 189–199, 2006.
- [Lesage2018] Pascal Lesage, Chris Mutel, Urs Schenker et Manuele Margni. *Uncertainty analysis in LCA using precalculated aggregated datasets*. The International Journal of Life Cycle Assessment, vol. 23, no. 11, pages 2248–2265, 2018.
- [Li2015] Bingdong Li, Jinlong Li, Ke Tang et Xin Yao. *Many-objective evolutionary algorithms: A survey*. ACM Computing Surveys (CSUR), vol. 48, no. 1, pages 1–35, 2015.
- [Li2018] Kaiwen Li, Rui Wang, Tao Zhang et Hisao Ishibuchi. *Evolutionary many-objective optimization: A comparative study of the state-of-the-art*. IEEE Access, vol. 6, pages 26194–26214, 2018.
- [Marler2004] R Timothy Marler et Jasbir S Arora. *Survey of multi-objective optimization methods for engineering*. Structural and multidisciplinary optimization, vol. 26, no. 6, pages 369–395, 2004.
- [Mastrucci2017] Alessio Mastrucci, Paula Perez-Lopez, Enrico Benetto, Ulrich Leopold et Isabelle Blanc. *Global sensitivity analysis as a support for the generation of simplified building stock energy models*. Energy and Buildings, vol. 149, pages 368–383, 2017.

-
- [Mehmeti2018] Andi Mehmeti, Athanasios Angelis-Dimakis, Carlos Boigues Muñoz, Marco Graziadio et Stephen J. McPhail. *Eco-thermodynamics of hydrogen production by high-temperature electrolysis using solid oxide cells*. Journal of Cleaner Production, vol. 199, pages 723–736, October 2018.
- [Moni2020] Sheikh Moniruzzaman Moni, Roksana Mahmud, Karen High et Michael Carbajales-Dale. *Life cycle assessment of emerging technologies: A review*. Journal of Industrial Ecology, vol. 24, no. 1, pages 52–63, February 2020.
- [Moore1959] Frederick T. Moore. *Economies of Scale: Some Statistical Evidence*. The Quarterly Journal of Economics, vol. 73, no. 2, page 232, May 1959.
- [Morbidoni2010] Alessandro Morbidoni, Marco Recchioni, Harald E Otto et Ferruccio Mandorli. *Enabling an efficient SLCA by interfacing selected PLM LCI parameters*. In Proceedings of TMCE, 2010.
- [Mutel2017] Chris Mutel. *Brightway: An open source framework for Life Cycle Assessment*. The Journal of Open Source Software, vol. 2, no. 12, page 236, April 2017.
- [Nadal2019] Amélia Nadal. *Impact of uncertainties on the techno-economic optimization of hybrid energy systems*. Theses, Communauté Université Grenoble Alpes, November 2019.
- [Nadal2020] Amélia Nadal, Alain Ruby, Cyril Bourasseau, Delphine Riu et Christophe Berenguer. *Accounting for techno-economic parameters uncertainties for robust design of remote microgrid*. International Journal of Electrical Power & Energy Systems, vol. 116, page 105531, March 2020.
- [Nagapurkar2019] Prashant Nagapurkar et Joseph D Smith. *Techno-economic optimization and environmental Life Cycle Assessment (LCA) of microgrids located in the US using genetic algorithm*. Energy Conversion and Management, vol. 181, pages 272–291, 2019.
- [Nazari-Heris2018] Morteza Nazari-Heris et Behnam Mohammadi-Ivatloo. *Application of Robust Optimization Method to Power System Problems*. In Classical and Recent Aspects of Power System Optimization, pages 19–32. Elsevier, 2018.
- [Ng2020] CY Ng, SS Lam, Samuel PM Choi et Kris MY Law. *Optimizing green design using ant colony-based approach*. The International Journal of Life Cycle Assessment, vol. 25, no. 3, pages 600–610, 2020.
- [Nordelöf2018] Anders Nordelöf et Anne-Marie Tillman. *A scalable life cycle inventory of an electrical automotive traction machine—Part II: manufacturing processes*. The International Journal of Life Cycle Assessment, vol. 23, no. 2, pages 295–313, February 2018.
- [Nugent2014] Daniel Nugent et Benjamin K Sovacool. *Assessing the lifecycle greenhouse gas emissions from solar PV and wind energy: A critical meta-survey*. Energy Policy, vol. 65, pages 229–244, 2014.
-

- [Nuss2014] Philip Nuss et Matthew J. Eckelman. *Life Cycle Assessment of Metals: A Scientific Synthesis*. PLoS ONE, vol. 9, no. 7, page e101298, July 2014.
- [O'Donoghue2014] Patrick R O'Donoghue, Garvin A Heath, Stacey L Dolan et Martin Vorum. *Life cycle greenhouse gas emissions of electricity generated from conventionally produced natural gas: systematic review and harmonization*. Journal of Industrial Ecology, vol. 18, no. 1, pages 125–144, 2014.
- [Ostad-Ahmad-Ghorabi2011] Hesamedin Ostad-Ahmad-Ghorabi et Daniel Collado-Ruiz. *Tool for the environmental assessment of cranes based on parameterization*. The International Journal of Life Cycle Assessment, vol. 16, no. 5, pages 392–400, June 2011.
- [Padey2013] Pierryves Padey, Robin Girard, Denis Le Boulch et Isabelle Blanc. *From LCAs to simplified models: a generic methodology applied to wind power electricity*. Environmental science & technology, vol. 47, no. 3, pages 1231–1238, 2013.
- [Pannier2018] Marie-Lise Pannier, Patrick Schalbart et Bruno Peuportier. *Comprehensive assessment of sensitivity analysis methods for the identification of influential factors in building life cycle assessment*. Journal of Cleaner Production, vol. 199, pages 466–480, October 2018.
- [Park2003] Ji-Hyung Park et Kwang-Kyu Seo. *Knowledge-based approximate life cycle assessment system in the collaborative design environment*. In 2003 EcoDesign 3rd International Symposium on Environmentally Conscious Design and Inverse Manufacturing, pages 499–503. IEEE, 2003.
- [Pascual-González2015] Janire Pascual-González, Carlos Pozo, Gonzalo Guillén-Gosálbez et Laureano Jiménez-Esteller. *Combined use of MILP and multi-linear regression to simplify LCA studies*. Computers & Chemical Engineering, vol. 82, pages 34–43, 2015.
- [Perez Gallardo2013] Jorge Raúl Perez Gallardo. *Ecodesign of large-scale photovoltaic (PV) systems with multi-objective optimization and Life-Cycle Assessment (LCA)*. Thèse de Doctorat, 2013.
- [Perez-Gallardo2018] J Raul Perez-Gallardo, Catherine Azzaro-Pantel et Stéphan Astier. *Combining multi-objective optimization, principal component analysis and multiple criteria decision making for ecodesign of photovoltaic grid-connected systems*. Sustainable Energy Technologies and Assessments, vol. 27, pages 94–101, 2018.
- [Pérez-López2018] Paula Pérez-López, Mahdokht Montazeri, Gumersindo Feijoo, María Teresa Moreira et Matthew J Eckelman. *Integrating uncertainties to the combined environmental and economic assessment of algal biorefineries: A Monte Carlo approach*. Science of the total environment, vol. 626, pages 762–775, 2018.
- [Pfenninger2016] Stefan Pfenninger et Iain Staffell. *Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data*. Energy, vol. 114, pages 1251–1265, 2016.

-
- [Piccinno2016] Fabiano Piccinno, Roland Hischer, Stefan Seeger et Claudia Som. *From laboratory to industrial scale: a scale-up framework for chemical processes in life cycle assessment studies*. Journal of Cleaner Production, vol. 135, pages 1085–1097, November 2016.
- [Pilenga2018] A Pilenga, G Tsotridis, European Commission et Joint Research Centre. EU harmonised terminology for low temperature water electrolysis for energy storage applications. 2018. OCLC: 1111109225.
- [Pirlo2019] G. Pirlo et S. Lolli. *Environmental impact of milk production from samples of organic and conventional farms in Lombardy (Italy)*. Journal of Cleaner Production, vol. 211, pages 962–971, February 2019.
- [Pizzol2017] Massimo Pizzol, Alexis Laurent, Serenella Sala, Bo Weidema, Francesca Verones et Christoph Koffler. *Normalisation and weighting in life cycle assessment: quo vadis?* The International Journal of Life Cycle Assessment, vol. 22, no. 6, pages 853–866, 2017.
- [Pozo2012] C Pozo, R Ruiz-Femenia, J Caballero, G Guillén-Gosálbez et L Jiménez. *On the use of Principal Component Analysis for reducing the number of environmental objectives in multi-objective optimization: Application to the design of chemical supply chains*. Chemical Engineering Science, vol. 69, no. 1, pages 146–158, 2012.
- [Ravindra2004] Khaiwal Ravindra, László Bencs et René Van Grieken. *Platinum group elements in the environment and their health risk*. Science of The Total Environment, vol. 318, no. 1-3, pages 1–43, January 2004.
- [Roches2010] Anne Roches, Thomas Nemecek, Gérard Gaillard, Katharina Plassmann, Sarah Sim, Henry King et Llorenç Milà i Canals. *MEX-ALCA: a modular method for the extrapolation of crop LCA*. The International Journal of Life Cycle Assessment, vol. 15, no. 8, pages 842–854, 2010.
- [Rubin2015] Edward S Rubin, Inês ML Azevedo, Paulina Jaramillo et Sonia Yeh. *A review of learning rates for electricity supply technologies*. Energy Policy, vol. 86, pages 198–218, 2015.
- [Sabio2012] N Sabio, A Kostin, G Guillén-Gosálbez et L Jiménez. *Holistic minimization of the life cycle environmental impact of hydrogen infrastructures using multi-objective optimization and principal component analysis*. International Journal of Hydrogen Energy, vol. 37, no. 6, pages 5385–5405, 2012.
- [Sacchi2019] Romain Sacchi, Romain Besseau, Paula Perez-Lopez et Isabelle Blanc. *Exploring technologically, temporally and geographically-sensitive life cycle inventories for wind turbines: A parameterized model for Denmark*. Renewable Energy, vol. 132, pages 1238–1250, 2019.
-

- [Saltelli2008] Andrea Saltelli, Marco Ratto, Terry Andres, Francesca Campolongo, Jessica Cariboni, Debora Gatelli, Michaela Saisana et Stefano Tarantola. *Global sensitivity analysis: the primer*. John Wiley & Sons, 2008.
- [Santos2019] Ruben Santos, António Aguiar Costa, José D Silvestre et Lincy Pyl. *Integration of LCA and LCC analysis within a BIM-based environment*. *Automation in Construction*, vol. 103, pages 127–149, 2019.
- [Santos2020] Rúben Santos, António Aguiar Costa, José D Silvestre et Lincy Pyl. *Development of a BIM-based Environmental and Economic Life Cycle Assessment tool*. *Journal of Cleaner Production*, page 121705, 2020.
- [Saxena2012] Dhish Kumar Saxena, João A Duro, Ashutosh Tiwari, Kalyanmoy Deb et Qingfu Zhang. *Objective reduction in many-objective optimization: Linear and nonlinear algorithms*. *IEEE Transactions on Evolutionary Computation*, vol. 17, no. 1, pages 77–99, 2012.
- [Schmidt Rivera2018] Ximena C. Schmidt Rivera, Evangelia Topriská, Maria Kolokotroni et Adisa Azapagic. *Environmental sustainability of renewable hydrogen in comparison with conventional cooking fuels*. *Journal of Cleaner Production*, vol. 196, pages 863–879, September 2018.
- [Schmidt2017] O. Schmidt, A. Gambhir, I. Staffell, A. Hawkes, J. Nelson et S. Few. *Future cost and performance of water electrolysis: An expert elicitation study*. *International Journal of Hydrogen Energy*, vol. 42, no. 52, pages 30470–30492, December 2017.
- [Schulz2012] Matthias Schulz, Michael D Short et Gregory M Peters. *A streamlined sustainability assessment tool for improved decision making in the urban water industry*. *Integrated Environmental Assessment and Management*, vol. 8, no. 1, pages 183–193, 2012.
- [Sharma2019] Hemant Sharma, Élise Monnier, Guillaume Mandil, Peggy Zwolinski et Stephane Colasson. *Comparison of environmental assessment methodology in Hybrid energy system simulation software*. *Procedia CIRP*, vol. 80, pages 221–227, 2019.
- [Sharma2020] Hemant Sharma, Guillaume Mandil, Peggy Zwolinski, Emmanuelle Cor, Hugo Mugnier et Elise Monnier. *Integration of life cycle assessment with energy simulation software for polymer exchange membrane (PEM) electrolysis*. *Procedia CIRP*, vol. 90, pages 176–181, 2020.
- [Simon2016] Bálint Simon, Krystyna Bachtin, Ali Kiliç, Ben Amor et Marcel Weil. *Proposal of a framework for scale-up life cycle inventory: A case of nanofibers for lithium iron phosphate cathode applications: Cradle-to-Gate LCA of Nanofiber Li-ion Battery Cathodes*. *Integrated Environmental Assessment and Management*, vol. 12, no. 3, pages 465–477, July 2016.
- [Sinha2014] Sunanda Sinha et S.S. Chandel. *Review of software tools for hybrid renewable energy systems*. vol. 32, pages 192–205, 2014.

-
- [Skansi2018] Sandro Skansi. *Introduction to deep learning: from logical calculus to artificial intelligence*. Springer, 2018.
- [Sobol'2010] I.M. Sobol' et S. Kucherenko. *A new derivative based importance criterion for groups of variables and its link with the global sensitivity indices*. *Computer Physics Communications*, vol. 181, no. 7, pages 1212–1217, July 2010.
- [Sobol'2009] I.M. Sobol' et S. Kucherenko. *Derivative based global sensitivity measures and their link with global sensitivity indices*. *Mathematics and Computers in Simulation*, vol. 79, no. 10, pages 3009–3017, June 2009.
- [Song2017] Runsheng Song, Arturo A Keller et Sangwon Suh. *Rapid life-cycle impact screening using artificial neural networks*. *Environmental science & technology*, vol. 51, no. 18, pages 10777–10785, 2017.
- [Staffell2016] Iain Staffell et Stefan Pfenninger. *Using bias-corrected reanalysis to simulate current and future wind power output*. *Energy*, vol. 114, pages 1224–1239, 2016.
- [Steinmann2014] Zoran JN Steinmann, Aranya Venkatesh, Mara Hauck, Aafke M Schipper, Ramkumar Karuppiah, Ian J Laurenzi et Mark AJ Huijbregts. *How to address data gaps in life cycle inventories: a case study on estimating CO₂ emissions from coal-fired electricity plants on a global scale*. *Environmental science & technology*, vol. 48, no. 9, pages 5282–5289, 2014.
- [Steinmann2016] Zoran JN Steinmann, Aafke M Schipper, Mara Hauck et Mark AJ Huijbregts. *How many environmental impact indicators are needed in the evaluation of product life cycles?* *Environmental Science & Technology*, vol. 50, no. 7, pages 3913–3919, 2016.
- [Sundaravaradan2011] Naren Sundaravaradan, Manish Marwah, Amip Shah et Naren Ramakrishnan. *Data mining approaches for life cycle assessment*. In *Proceedings of the 2011 IEEE international symposium on sustainable systems and technology*, pages 1–6. IEEE, 2011.
- [Tasala Gradin2020] Katja Tasala Gradin. *Simplified Life Cycle Assessment Approaches and Potential Impact Shifts*. Thèse de Doctorat, Kungliga Tekniska högskolan, 2020.
- [Taylor1990] Richard Taylor. *Interpretation of the correlation coefficient: a basic review*. *Journal of diagnostic medical sonography*, vol. 6, no. 1, pages 35–39, 1990.
- [Tecchio2016] Paolo Tecchio, Pierluigi Freni, Bruno De Benedetti et Françoise Fenouillot. *Ex-ante Life Cycle Assessment approach developed for a case study on bio-based polybutylene succinate*. *Journal of Cleaner Production*, vol. 112, pages 316–325, January 2016.
- [Thomassen2019] Gwenny Thomassen, Miet Van Dael, Steven Van Passel et Fengqi You. *How to assess the potential of emerging green technologies? Towards a prospective environmental and techno-economic assessment framework*. *Green Chemistry*, vol. 21, no. 18, pages 4868–4886, 2019.
-

- [Thomassen2020] Gwenny Thomassen, Steven Van Passel et Jo Dewulf. *A review on learning effects in prospective technology assessment*. *Renewable and Sustainable Energy Reviews*, vol. 130, page 109937, 2020.
- [Tietze2020] Ingela Tietze, Lukas Lazar, Heidi Hottenroth et Steffen Lewerenz. *LAEND: A Model for Multi-Objective Investment Optimisation of Residential Quarters Considering Costs and Environmental Impacts*. *Energies*, vol. 13, no. 3, page 614, 2020.
- [Touzani2014] Samir Touzani et Daniel Busby. *Screening Method Using the Derivative-based Global Sensitivity Indices with Application to Reservoir Simulator*. *Oil & Gas Science and Technology – Revue d'IFP Energies nouvelles*, vol. 69, no. 4, pages 619–632, 2014.
- [Turconi2013] Roberto Turconi, Alessio Boldrin et Thomas Astrup. *Life cycle assessment (LCA) of electricity generation technologies: Overview, comparability and limitations*. *Renewable and sustainable energy reviews*, vol. 28, pages 555–565, 2013.
- [Valente2017] Antonio Valente, Diego Iribarren et Javier Dufour. *Harmonised life-cycle global warming impact of renewable hydrogen*. *Journal of Cleaner Production*, vol. 149, pages 762–772, April 2017.
- [van der Hulst2020] Mitchell K van der Hulst, Mark AJ Huijbregts, Niels van Loon, Mirjam Theelen, Lucinda Kootstra, Joseph D Bergesen et Mara Hauck. *A systematic approach to assess the environmental impact of emerging technologies: A case study for the GHG footprint of CIGS solar photovoltaic laminate*. *Journal of Industrial Ecology*, 2020.
- [van Kalkereren2013] Henri A. van Kalkereren, Anneloes L. Blom, Floris P. J. T. Rutjes et Mark A. J. Huijbregts. *On the usefulness of life cycle assessment in early chemical methodology development: the case of organophosphorus-catalyzed Appel and Wittig reactions*. *Green Chemistry*, vol. 15, no. 5, page 1255, 2013.
- [Vandepaer2020] Laurent Vandepaer, Evangelos Panos, Christian Bauer et Ben Amor. *Energy System Pathways with Low Environmental Impacts and Limited Costs: Minimizing Climate Change Impacts Produces Environmental Cobenefits and Challenges in Toxicity and Metal Depletion Categories*. *Environmental Science & Technology*, vol. 54, no. 8, pages 5081–5092, 2020.
- [Vázquez2018] Daniel Vázquez, Rubén Ruiz-Femenia, Laureano Jiménez et José A Caballero. *MILP models for objective reduction in multi-objective optimization: Error measurement considerations and non-redundancy ratio*. *Computers & Chemical Engineering*, vol. 115, pages 323–332, 2018.
- [Verghese2010] Karli L. Verghese, Ralph Horne et Andrew Carre. *PIQET: the design and development of an online 'streamlined' LCA tool for sustainable packaging design decision support*. *The International Journal of Life Cycle Assessment*, vol. 15, no. 6, pages 608–620, July 2010.

-
- [Wei2015] Wei Wei, Pyrene Larrey-Lassalle, Thierry Faure, Nicolas Dumoulin, Philippe Roux et Jean-Denis Mathias. *How to Conduct a Proper Sensitivity Analysis in Life Cycle Assessment: Taking into Account Correlations within LCI Data and Interactions within the LCA Calculation Model*. *Environmental Science & Technology*, vol. 49, no. 1, pages 377–385, January 2015.
- [Weidema2015] Bo P Weidema. *Comparing three life cycle impact assessment methods from an endpoint perspective*. *Journal of Industrial Ecology*, vol. 19, no. 1, pages 20–26, 2015.
- [Wernet2008] Gregor Wernet, Stefanie Hellweg, Ulrich Fischer, Stavros Papadokostantakis et Konrad Hungerbuhler. *Molecular-structure-based models of chemical inventories using neural networks*. *Environmental science & technology*, vol. 42, no. 17, pages 6717–6722, 2008.
- [Wernet2012] Gregor Wernet, Stefanie Hellweg et Konrad Hungerbuhler. *A tiered approach to estimate inventory data and impacts of chemical products and mixtures*. *The International Journal of Life Cycle Assessment*, vol. 17, no. 6, pages 720–728, July 2012.
- [Wernet2016] Gregor Wernet, Christian Bauer, Bernhard Steubing, Jürgen Reinhard, Emilia Moreno-Ruiz et Bo Weidema. *The ecoinvent database version 3 (part I): overview and methodology*. *The International Journal of Life Cycle Assessment*, vol. 21, no. 9, pages 1218–1230, September 2016.
- [While2011] Lyndon While, Lucas Bradstreet et Luigi Barone. *A fast way of calculating exact hypervolumes*. *IEEE Transactions on Evolutionary Computation*, vol. 16, no. 1, pages 86–95, 2011.
- [Whitaker2012] Michael Whitaker, Garvin A Heath, Patrick O'Donoghue et Martin Vorum. *Life cycle greenhouse gas emissions of coal-fired electricity generation: Systematic review and harmonization*. *Journal of Industrial Ecology*, vol. 16, pages S53–S72, 2012.
- [Wild-Scholten2013] M.J. Wild-Scholten. *Energy payback time and carbon footprint of commercial photovoltaic systems*. *Solar Energy Materials and Solar Cells*, vol. 119, pages 296–305, December 2013.
- [Wolfgang2011] Ertel Wolfgang. *Introduction to artificial intelligence*. Springer, 2011.
- [Yeh2012] Sonia Yeh et Edward S Rubin. *A review of uncertainties in technology experience curves*. *Energy Economics*, vol. 34, no. 3, pages 762–771, 2012.
- [Yi2011] Sora Yi, Kiyo H Kurisu et Keisuke Hanaki. *Life cycle impact assessment and interpretation of municipal solid waste management scenarios based on the midpoint and endpoint approaches*. *The International Journal of Life Cycle Assessment*, vol. 16, no. 7, pages 652–668, 2011.
-

- [Yuan2017] Yuan Yuan, Yew-Soon Ong, Abhishek Gupta et Hua Xu. *Objective reduction in many-objective optimization: evolutionary multiobjective approaches and comprehensive analysis*. IEEE Transactions on Evolutionary Computation, vol. 22, no. 2, pages 189–210, 2017.
- [Yue2016] Dajun Yue, Shyama Pandya et Fengqi You. *Integrating hybrid life cycle assessment with multiobjective optimization: a modeling framework*. Environmental science & technology, vol. 50, no. 3, pages 1501–1509, 2016.
- [Zaiontz2017] Charles Zaiontz. *Real statistics using Excel. 2017*. Computer software] <https://www.real-statistics.com/free-download/>(accessed 20 Oct 2020), 2017.
- [Zhang2015] Di Zhang, Sara Evangelisti, Paola Lettieri et Lazaros G Papageorgiou. *Optimal design of CHP-based microgrids: Multiobjective optimisation and life cycle assessment*. Energy, vol. 85, pages 181–193, 2015.
- [Zhang2017] Xiaojin Zhang, Christian Bauer, Christopher L. Mutel et Kathrin Volkart. *Life Cycle Assessment of Power-to-Gas: Approaches, system variations and their environmental implications*. Applied Energy, vol. 190, pages 326–338, March 2017.
- [Zhao2018] Guangling Zhao et Eva Ravn Nielsen. *Environmental Impact Study of BIG HIT*, 2018.
- [Zhou2015] Qifeng Zhou, Hao Zhou, Yimin Zhu et Tao Li. *Data-driven solutions for building environmental impact assessment*. In Proceedings of the 2015 IEEE 9th International Conference on Semantic Computing (IEEE ICSC 2015), pages 316–319. IEEE, 2015.
- [Zimmermann2019] R Kjær Zimmermann, Kai Kanafani, F Nygaard Rasmussen et Harpa Birgisdóttir. *Early Design Stage Building LCA using the LCAByg tool: Comparing Cases for Early Stage and Detailed LCA Approaches*. In IOP Conference Series: Earth and Environmental Science, volume 323, page 012118. IOP Publishing, 2019.
- [Zitzler2001] Eckart Zitzler, Marco Laumanns et Lothar Thiele. *SPEA2: Improving the strength Pareto evolutionary algorithm*. TIK-report, vol. 103, 2001.
- [Zitzler2003] Eckart Zitzler, Lothar Thiele, Marco Laumanns, Carlos M Fonseca et Viviane Grunert Da Fonseca. *Performance assessment of multi-objective optimizers: An analysis and review*. IEEE Transactions on evolutionary computation, vol. 7, no. 2, pages 117–132, 2003.
- [Zitzler2007] Eckart Zitzler, Dimo Brockhoff et Lothar Thiele. *The hypervolume indicator revisited: On the design of Pareto-compliant indicators via weighted integration*. In International Conference on Evolutionary Multi-Criterion Optimization, pages 862–876. Springer, 2007.

List of Figures

1	Breakdown of world greenhouse gas emissions in 2016, updated from [Herzog2009].	1
1.1	An overview of the energy simulation software used in this thesis - Odyssey. . . .	7
1.2	Crossover and mutation operations, illustration from [Perez Gallardo2013]. . . .	9
1.3	SPEA2 operating scheme as described in [Zitzler2001].	10
1.4	Case study of hydrogen production from PEM water electrolysis.	11
1.5	Inputs time series: hydrogen demand (kg/h) and electricity price (€/MWh). . . .	11
1.6	Pareto front of the results for the dimensioning problem	12
1.7	Output time series for case no. 3: electrolyser power (MW) and Hydrogen quantity in the tank (kg)	13
1.8	(a) Life cycle assessment framework adapted from ILCD handbook [European Commission2010]; (b) A conceptual representation of inventory building and its use for impact calculation step of an LCA.	16
1.9	User preferences from the environmental modelling rated from 1-10 with 10 being very important. Blue dots represent the user rating and the red cross represents the average.	17
1.10	Representation of the case study in Odyssey	19
1.11	Pareto front of the results for the dimensioning problem	20
1.12	A simplified illustration of the research approach followed in this thesis	22
2.1	Breakdown of climate change potential according to different life cycle stages for electricity production from wind turbines, solar PV, conventional natural gas and coal fired plants [Sharma2019].	26
2.2	The search string used to identify documents in Scopus.	29
2.3	Statistics of the total 165 references analysed. (<i>left</i>) Year-wise distribution; (<i>right</i>) Sector-wise distribution.	30
2.4	(<i>Left</i>) Simplified illustration of an artificial neural network; (<i>Right</i>) An artificial neuron, simplified., illustration adapted from [Skansi2018]	33
2.5	(<i>Left</i>) Objects for clustering, representation in a two dimensional data (<i>Right</i>) Clustering of objects [Bramer2007].	35
2.6	NSGA-III operating scheme in this thesis.	40
2.7	Reference point distribution illustrated for 3 dimensions. (<i>Left</i>) for 10 axis partitions. (<i>Right</i>) using a layering approach when the number of objectives increase.	41
2.8	Illustration of the association operation of individuals (orange points) with reference points (in blue) [Deb2013].	42
2.9	General objective reduction scheme used for optimisation including LCA indicators	45
2.10	An illustration of the methodology to select indicators using PCA heuristics by [Deb2005, Saxena2012].	45
2.11	An example to demonstrate dominance-structure based reduction technique. (a) Three original objectives (b) Objective f_3 discarded (b) Objective f_2 discarded . Illustration from [Vázquez2018].	47

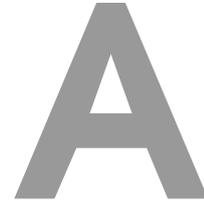
3.1	A schematic representation of the proposed framework. The steps that are described in this chapter are outlined in black.	55
3.2	Proposed scaling approach for dimensioning inventories of components. Adapted from [Gerber2011].	59
3.3	Representing the parameterized model and disaggregating the impact equation in terms of impact contributions and energy flow.	62
3.4	Interaction of sensitivity analysis package with energy simulations and LCA modelling.	65
3.5	Illustration of the case study and the LCA system boundaries	67
3.6	Description of the parameterized model for the PEM water electrolysis case study. New parameters added, that are relevant to LCA are illustrated in green.	71
3.7	Sensitivity indices (γ) obtained for the parameters for each of the selected mid-point indicators	72
3.8	Results obtained for the LCA indicators for the three cases.	75
3.9	Boxplots of distribution of impacts obtained by propagating the PEM electrolyser system efficiency and electricity mix using quasi-monte carlo. Maximum and minimum values obtained by propagating extreme points of efficiency and electricity mix using LSA	76
4.1	Simplified representation of climate change impact pathway from elementary flows to the areas of protection [Hauschild2015].	81
4.2	Impact indicators at endpoint and midpoints along with their impact pathways in impact assessment method ReCiPe 2016 [Huijbregts2016].	81
4.3	An illustration of the methodology to select impact indicators using a visual survey	84
4.4	Classification of the indicators according to the visual survey (steps 1-4) applied to a fictional case study of designing a bike	85
4.5	Results of impact indicators of the 20 design alternatives	86
4.6	Contribution analysis of the 5 solutions in terms of the two LCA indicators.	86
4.7	Scree plot of the principal components	88
4.8	LCA indicators plotted in terms of the two principal components.	89
5.1	Representation of the case study in Odyssey.	94
5.2	Production data of solar PV and Wind for the first nine days.	95
5.3	Computational structure for search of solutions for including technical, economic and LCA indicators	98
5.4	Comparison scheme for testing performance for different optimisation cases	99
5.5	Hypervolume calculation principle adapted from [Zitzler2003].	99
5.6	Indicator selection procedure using visual survey for this case study	101
5.7	(Left) Scree plot of principal components are their cumulative explained variability. (Right) First two principal components plotted against each other to reveal correlations between indicators.	104
5.8	Hypervolume indicator for different simulation runs.	106
5.9	Distribution of the pareto front obtained using 14 and 7 objectives shown in four figures of 2-dimensional plots	107
5.10	Distribution of the pareto front obtained using 7 and 4 objectives shown in four figures of 2-dimensional plots	108
5.11	Distribution of the pareto front obtained using 7 and 4 objectives shown for two, 2D plots	109
5.12	Standardised minimum value of objectives vs. number of evaluations for different simulation runs.	109
5.13	Proposed approach to select and visualize solutions.	112

5.14	Replacing the dominated population member by a similar individual in the pareto front using individual or fitness similarities.	114
5.15	Population spread in the pareto.	114
5.16	Final set of solutions provided in a visual tool.	115
5.17	The correlations between 4 indicators retained and the other 10 discarded in the final set of 120 solutions	116
5.18	Comparison of tech-eco optimisation and with LCA. Black point is solution no. 20 in figure 5.16.	118
5.19	Illustration of the proposed approach to include LCA indicators in the assessments of hybrid energy systems	123
A.1	PV modifications	XXXI
C.1	Hypervolume indicator for different number of objectives with NSGA-III with extended number of evaluations	XLI
C.2	Hypervolume indicator for different number of objectives with NSGA-III with number of generations	XLII
C.3	Distribution of the pareto front obtained using 7 and 14 objectives shown in four figures of 2-dimensional plots	XLII
C.4	Distribution of the pareto front obtained using 7 and 14 objectives shown in four figures of 2-dimensional plots	XLIII
C.5	Distribution of the pareto front obtained using 7 and 4 objectives shown in four figures of 2-dimensional plots	XLIII
C.6	Distribution of the pareto front obtained using 7 and 4 objectives shown in four figures of 2-dimensional plots	XLIV
C.7	Distribution of the pareto front obtained using 7 and 2 objectives shown in four figures of 2-dimensional plots	XLIV
C.8	Distribution of the pareto front obtained using 7 and 2 objectives shown in four figures of 2-dimensional plots	XLV
C.9	Evaluations vs. standardised minimum point of indicators in the pareto front	XLVI

List of Tables

1.1	Summary of inputs to the optimisation	12
1.2	Description of the four selected cases from the pareto front	12
1.3	Search range provided to the genetic algorithm.	19
1.4	Summary of the obtained 156 system configurations obtained at the end of the search	19
2.1	Summary of the environmental analysis in the reviewed software [Sharma2019].	25
2.2	Summary of the environmental analysis in the reviewed software [Sharma2019].	27
2.3	Fictitious train data to demonstrate naive Bayes classifier [Skansi2018].	36
2.4	Recommendations on the use of the LCA modelling methods and relevance of their results	37
3.1	Input ranges to the DGSM sensitivity analysis	72
3.2	Importance indices (Im) for 3 arbitrarily selected foreground processes from the use phase and the stack inventory. Full results in annex table A.6.	73
3.3	Comparison of input parameters for the three cases	74
4.1	Correlation matrix of 8 impact indicators for the 20 bike alternatives.	87
4.2	PCA results for the 8 impact indicators. Selected indicators are highlighted in bold.	88
4.3	Comparison of indicators selected by visual survey and PCA	89
5.1	Summary of inputs to the integrated techno-economic-LCA model	96
5.2	PCA results for the 14 indicators. Selected indicators are highlighted in bold.	103
5.3	Indicators retained by PCA during the five simulation runs	104
5.4	Reference points, population size and indicators retained for the different simulation runs.	105
5.5	Estimation of computational time using multi-processing for different number of indicators.	111
5.6	Comparison of indicators selected using the visual technique and using a particular setting in PCA.	112
5.7	Correlation coefficients for retained indicators with the discarded ones (except freshwater ecotoxicity). Full results in table A.13.	116
A.1	Stack inventory for the three cases, data from [Bareiß2019], electricity requirement from [Evangelisti2017]	XXIX
A.2	Balance of plant inventory, data from [Zhang2017] and [Bareiß2019]	XXX
A.3	Primary production activities	XXX
A.4	Disposal activities	XXX
A.5	Recycling activities	XXXII
A.6	Importance indices for foreground processes from the use phase and the stack inventory.	XXXII

A.7 Electricity mix utilised in the case study from ecoinvent 3.6 cutoff and for 2030 adapted from ADEME	XXXIII
A.8 Current allocation for platinum group metals in ecoinvent v3.6 and updated allocation using factors.	XXXIV
A.9 PV inventory activities	XXXIV
A.10 Absolute impact values from the three cases	XXXIV
A.11 Name of the three materials from ecoinvent 3.6 and absolute impact per kg for the design of bike	XXXV
A.12 Correlation matrix after 5 generations for indicator reduction using PCA, for one simulation run.	XXXV
A.13 Correlation matrix for indicator reduction using PCA obtained for final display of solutions	XXXVI



Annex 1 : Supplementary data

Table A.1: Stack inventory for the three cases, data from [Bareiß2019], electricity requirement from [Evangelisti2017]

Activity	Base case	Only imp. data adjusted	Ideal case
market for titanium, primary' (kilogram, GLO, None)	528.0	528.0	37.0
market for aluminium, primary, ingot' (kilogram, RoW, None)	27.0	27.0	54.0
market for steel, chromium steel 18/8, hot rolled' (kilogram, GLO, None)	100.0	100.0	40.0
market for copper, cathode' (kilogram, GLO, None)	4.5	4.5	9.0
market for sulfuric acid' (kilogram, RER, None)	13.6	13.6	1.7
market for tetrafluoroethylene film, on glass' (kilogram, GLO, None)	2.88	2.88	0.36
market for activated carbon, granular' (kilogram, GLO, None)	9	9	4.5
market for iridium' (kilogram, GLO, None)	0	0.037	0.037
market for platinum' (kilogram, GLO, None)	0.825	0.01	0.01
electricity, high voltage, production mix' (kilowatt hour, FR, None)	5360.0	5360.0	5360.0

Table A.2: Balance of plant inventory, data from [Zhang2017] and [Bareiß2019]

Activity	Amount
market for steel, unalloyed' (kilogram, GLO, None)	13700.0
market for copper' (kilogram, GLO, None)	5900.0
market for steel, chromium steel 18/8, hot rolled' (kilogram, GLO, None)	1100.0
market for aluminium, primary, ingot' (kilogram, RoW, None)	1000.0
market for polyethylene, high density, granulate, recycled' (kilogram, CH, None)	300.0
market for concrete, 30-32MPa' (cubic meter, RoW, None)	2.113
market for display, liquid crystal, 17 inches' (unit, GLO, None)	1.0

Table A.3: Primary production activities

Primary production activities
'copper production, primary' (kilogram, RER, None)
'pig iron production' (kilogram, RoW, None)
'aluminium production, primary, ingot' (kilogram, IAI Area, Russia & RER w/o EU27 & EFTA, None)
'steel production, converter, chromium steel 18/8' (kilogram, RER, None)
'primary lead production from concentrate' (kilogram, GLO, None)
'polyethylene production, high density, granulate' (kilogram, RER, None)
'steel production, converter, low-alloyed' (kilogram, RER, None)

Table A.4: Disposal activities

Disposal activities
'market for waste polyvinylchloride' (kilogram, FR, None)
'market for waste polypropylene' (kilogram, FR, None)
'market for waste concrete' (kilogram, RoW, None)
'market for waste mineral oil' (kilogram, CH, None)
'market for waste reinforcement steel' (kilogram, CH, None)
'market for waste rubber, unspecified' (kilogram, CH, None)
'market for waste aluminium' (kilogram, GLO, None)
'market for process-specific burdens, inert material landfill' (kilogram, RoW, None)

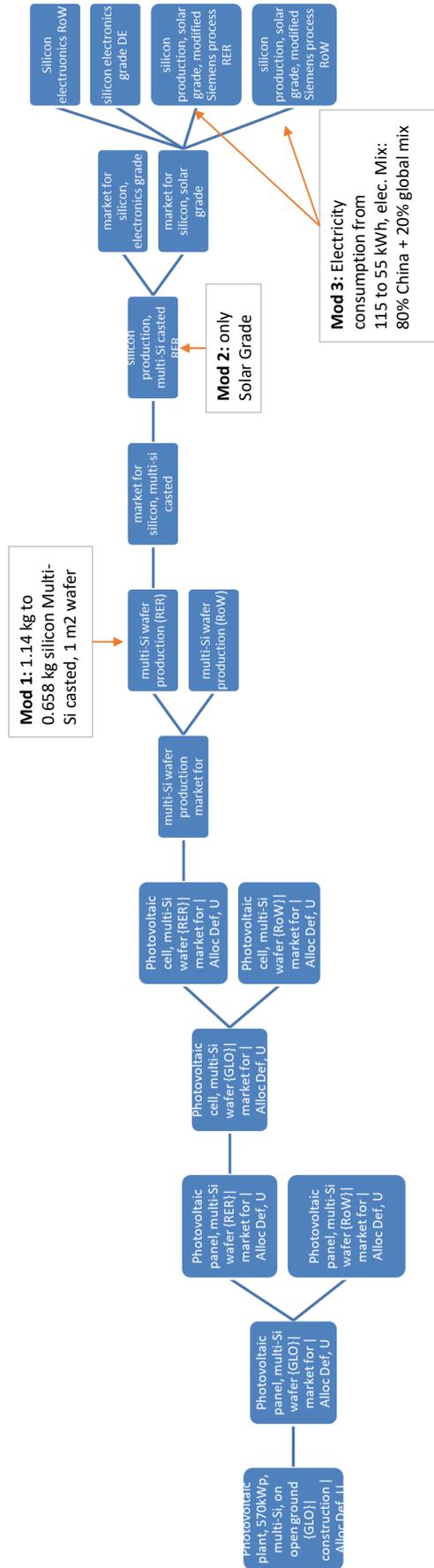


Figure A.1: PV modifications

Table A.5: Recycling activities

Recycling activities
'treatment of copper scrap by electrolytic refining' (kilogram, RER, None)
'market for iron scrap, sorted, pressed' (kilogram, GLO, None)
'treatment of aluminium scrap, new, at remelter' (kilogram, RER, None)
'treatment of aluminium scrap, post-consumer, prepared for recycling, at remelter' (kilogram, RER, None)
'treatment of scrap lead acid battery, remelting' (kilogram, RER, None)
'steel production, electric, chromium steel 18/8' (kilogram, RER, None)
'polyethylene production, high density, granulate, recycled' (kilogram, CH, None)
'treatment of waste polyethylene, for recycling, unsorted, sorting' (kilogram, CH, None)
'steel production, electric, low-alloyed' (kilogram, RER, None)

Table A.6: Importance indices for foreground processes from the use phase and the stack inventory.

Foreground processes	Classification	Importance indices (Im)						
		Climate change	Ozone Layer Depletion	Terrestrial acidification	Particulate matter formation	Photochemical oxidant formation	Human toxicity cancer	Mineral resources use
market for water, deionised	Use phase	0.0007	0.0007	0.0018	0.0015	0.0010	0.0019	0.0041
electricity, high voltage, production mix	Use phase	1.2124	1.2136	0.7999	1.0517	1.0660	1.1216	1.1959
market for titanium, primary	Manufacturing, PEM stack	0.0045	0.0001	0.0440	0.0406	0.0379	0.0134	0.0024
market for aluminium, primary, ingot	Manufacturing, PEM stack	0.0001	0.0000	0.0018	0.0015	0.0011	0.0023	0.0000
market for steel, chromium steel 18/8	Manufacturing, PEM stack	0.0001	0.0000	0.0015	0.0020	0.0012	0.0031	0.0003
market for copper, cathode	Manufacturing, PEM stack	0.0000	0.0000	0.0013	0.0006	0.0003	0.0003	0.0001
market for sulfuric acid	Manufacturing, PEM stack	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000
market for tetrafluoroethylene	Manufacturing, PEM stack	0.0001	0.0010	0.0001	0.0001	0.0001	0.0000	0.0000
market for activated carbon, granular	Manufacturing, PEM stack	0.0000	0.0000	0.0001	0.0001	0.0001	0.0000	0.0000
market for platinum	Manufacturing, PEM stack	0.0172	0.0004	2.0390	0.7944	0.7957	0.1638	0.0128
electricity, high voltage, production mix	Manufacturing, PEM stack	0.0001	0.0000	0.0008	0.0006	0.0006	0.0003	0.0001

Table A.7: Electricity mix utilised in the case study from ecoinvent 3.6 cutoff and for 2030 adapted from ADEME

Electricity mix activities from Ecoinvent 3.6	Electricity mix high voltage, France, v3.6.	Electricity mix France 2030, adapted using [ADEME2014]
'electricity production, deep geothermal' (kilowatt hour, FR, None)	0,00%	0,30%
'electricity production, hard coal' (kilowatt hour, FR, None)	1,49%	0,00%
'electricity production, hydro, pumped storage' (kilowatt hour, FR, None)	0,95%	0,89%
'electricity production, hydro, reservoir, alpine region' (kilowatt hour, FR, None)	1,87%	1,75%
'electricity production, hydro, run-of-river' (kilowatt hour, FR, None)	9,79%	9,19%
'electricity production, natural gas, combined cycle power plant' (kilowatt hour, FR, None)	3,62%	3,75%
'electricity production, natural gas, conventional power plant' (kilowatt hour, FR, None)	0,70%	0,72%
'electricity production, nuclear, pressure water reactor' (kilowatt hour, FR, None)	75,17%	48,00%
'electricity production, oil' (kilowatt hour, FR, None)	0,32%	0,00%
'electricity production, wind, 1-3MW turbine, offshore' (kilowatt hour, FR, None)	0,00%	11,22%
'electricity production, wind, 1-3MW turbine, onshore' (kilowatt hour, FR, None)	3,72%	11,22%
'electricity production, wind, <1MW turbine, onshore' (kilowatt hour, FR, None)	0,24%	0,00%
'electricity production, wind, >3MW turbine, onshore' (kilowatt hour, FR, None)	0,02%	0,00%
'heat and power co-generation, biogas, gas engine' (kilowatt hour, FR, None)	0,09%	1,83%
'heat and power co-generation, hard coal' (kilowatt hour, RoW, None)	0,00%	0,00%
'heat and power co-generation, natural gas, conventional power plant, 100MW electrical' (kilowatt hour, FR, None)	1,28%	1,33%
'heat and power co-generation, oil' (kilowatt hour, FR, None)	0,04%	0,00%
'heat and power co-generation, wood chips, 6667 kW, state-of-the-art 2014' (kilowatt hour, FR, None)	0,36%	0,61%
'treatment of blast furnace gas, in power plant' (kilowatt hour, FR, None)	0,29%	0,00%
'treatment of coal gas, in power plant' (kilowatt hour, FR, None)	0,05%	0,00%
electricity, from municipal waste incineration to generic market for electricity, medium voltage' (kilowatt hour, FR, None)	0,00%	0,61%
electricity production, photovoltaic, 570kWp open ground installation, multi-Si' (kilowatt hour, FR, None)	0,00%	9,15%

Table A.8: Current allocation for platinum group metals in ecoinvent v3.6 and updated allocation using factors.

Metals	Current allocation ecoinvent v3.6	Updated allocation
	[Wernet2016]	using [Nuss2014]
Nickel	11%	10,55%
Platinum	63%	62,02%
Copper	4%	3,65%
Palladium	10%	10,24%
Rhodium	11%	11,23%
Gold	1%	1,39%
Ruthenium	0%	0,41%
Osmium	0%	0,08%
Iridium	0%	0,43%

Table A.9: PV inventory activities

Activity	Location	Amount	Unit
market for diesel, burned in building machine	GLO	7673	megajoule
market for photovoltaics, electric installation for 570kWp module, open ground	GLO	1	unit
market for inverter, 500kW	GLO	3.126	unit
market for photovoltaic mounting system, for 570kWp open ground module	GLO	4273.5	square meter
market group for electricity, low voltage	GLO	36.033	kilowatt hour
photovoltaic panel production, multi-Si wafer updated LSED	RER	4401.7	square meter

Table A.10: Absolute impact values from the three cases

Case	Climate change (kg CO ₂ -eq)	Ozone depletion (kg CFC-11-eq 10 ⁻⁶)	Terrestrial acidification (kg SO ₂ -eq 10 ⁻²)	Particulate matter formation (kg PM _{2.5} -eq 10 ⁻⁶)	Photochemical Oxidant Formation (kg NMVOC-eq 10 ⁻³)	Human Toxicity cancer (CTUh 10 ⁻⁷)	Mineral resource use (100 kg deprived)
Base case	4.698	3.931	1.793	2.087	12.996	2.548	1.596
Only important parameters adjusted	3.584	2.651	1.253	1.694	9.182	2.952	1.852
Ideal case	3.568	2.642	1.243	1.679	9.103	2.965	1.845

Table A.11: Name of the three materials from ecoinvent 3.6 and absolute impact per kg for the design of bike

Activity name	Climate change (kg CO ₂ -eq)	Fossil and nuclear energy use (MJ deprived)	Freshwater ecotoxicity (CTUe)	Freshwater eutrophication (kg PO ₄ P-lim eq)	Human toxicity cancer (CTUh)	Marine eutrophication (kg N N-lim eq)	Mineral resources use (kg deprived)	Terrestrial acidification (kg SO ₂ eq)
aluminium production, primary, ingot RoW kilogram	18.428	193.292	460501.000	1.98E-02	5.18E-06	1.18E-03	0.023	0.091
glass fibre production RER kilogram	2.039	36.918	24422.516	2.47E-03	1.43E-07	3.33E-04	0.295	0.012
tin production RoW kilogram	10.086	144.557	3891383.069	7.97E-02	2.90E-06	2.10E-03	0.113	0.089

Table A.12: Correlation matrix after 5 generations for indicator reduction using PCA, for one simulation run.

	Cost	FS	CC	Min	W	FWA	TER	FW eco	PM	POF	HT can	HT non can	IO	Oz
Cost	1.00	-0.68	0.55	0.95	-0.45	0.69	0.69	0.68	0.77	0.77	0.95	0.75	-0.71	-0.65
FS	-0.68	1.00	-0.04	-0.72	0.92	-0.30	-0.30	-0.43	-0.41	-0.43	-0.76	-0.34	1.00	1.00
CC	0.55	-0.04	1.00	0.63	0.35	0.96	0.96	0.89	0.93	0.92	0.61	0.95	-0.11	0.03
Min	0.95	-0.72	0.63	1.00	-0.46	0.77	0.77	0.77	0.85	0.86	1.00	0.81	-0.77	-0.69
W	-0.45	0.92	0.35	-0.46	1.00	0.09	0.09	-0.05	-0.03	-0.05	-0.49	0.05	0.89	0.95
FWA	0.69	-0.30	0.96	0.77	0.09	1.00	1.00	0.98	0.99	0.99	0.77	1.00	-0.38	-0.23
TER	0.69	-0.30	0.96	0.77	0.09	1.00	1.00	0.98	0.99	0.99	0.77	1.00	-0.38	-0.23
FW eco	0.68	-0.43	0.89	0.77	-0.05	0.98	0.98	1.00	0.97	0.98	0.78	0.97	-0.50	-0.36
PM	0.77	-0.41	0.93	0.85	-0.03	0.99	0.99	0.97	1.00	1.00	0.85	1.00	-0.48	-0.34
POF	0.77	-0.43	0.92	0.86	-0.05	0.99	0.99	0.98	1.00	1.00	0.86	0.99	-0.50	-0.36
HT can	0.95	-0.76	0.61	1.00	-0.49	0.77	0.77	0.78	0.85	0.86	1.00	0.81	-0.80	-0.72
HT non can	0.75	-0.34	0.95	0.81	0.05	1.00	1.00	0.97	1.00	0.99	0.81	1.00	-0.41	-0.27
IO	-0.71	1.00	-0.11	-0.77	0.89	-0.38	-0.38	-0.50	-0.48	-0.50	-0.80	-0.41	1.00	0.99
Oz	-0.65	1.00	0.03	-0.69	0.95	-0.23	-0.23	-0.36	-0.34	-0.36	-0.72	-0.27	0.99	1.00

Table A.13: Correlation matrix for indicator reduction using PCA obtained for final display of solutions

	cost	FS	CC	Min	W	FWA	TER	FW eco	PM	POF	HT can	HT non can	IO	Oz
cost	1.00	-0.83	-0.39	0.98	-0.79	0.24	0.23	0.57	0.57	0.64	1.00	0.88	-0.84	-0.82
FS	-0.83	1.00	0.80	-0.76	1.00	0.20	0.20	-0.56	-0.14	-0.23	-0.84	-0.53	1.00	1.00
CC	-0.39	0.80	1.00	-0.27	0.85	0.73	0.73	-0.09	0.47	0.39	-0.40	0.07	0.79	0.81
Min	0.98	-0.76	-0.27	1.00	-0.71	0.35	0.35	0.58	0.66	0.73	0.99	0.93	-0.77	-0.75
W	-0.79	1.00	0.85	-0.71	1.00	0.28	0.28	-0.50	-0.06	-0.15	-0.80	-0.46	0.99	1.00
FWA	0.24	0.20	0.73	0.35	0.28	1.00	1.00	0.56	0.94	0.90	0.23	0.66	0.18	0.21
TER	0.23	0.20	0.73	0.35	0.28	1.00	1.00	0.56	0.93	0.90	0.23	0.66	0.18	0.21
FW eco	0.57	-0.56	-0.09	0.58	-0.50	0.56	0.56	1.00	0.69	0.72	0.58	0.67	-0.57	-0.55
PM	0.57	-0.14	0.47	0.66	-0.06	0.94	0.93	0.69	1.00	0.99	0.56	0.88	-0.16	-0.13
POF	0.64	-0.23	0.39	0.73	-0.15	0.90	0.90	0.72	0.99	1.00	0.64	0.92	-0.25	-0.22
HT can	1.00	-0.84	-0.40	0.99	-0.80	0.23	0.23	0.58	0.56	0.64	1.00	0.88	-0.85	-0.84
HT non can	0.88	-0.53	0.07	0.93	-0.46	0.66	0.66	0.67	0.88	0.92	0.88	1.00	-0.54	-0.52
IO	-0.84	1.00	0.79	-0.77	0.99	0.18	0.18	-0.57	-0.16	-0.25	-0.85	-0.54	1.00	1.00
Oz	-0.82	1.00	0.81	-0.75	1.00	0.21	0.21	-0.55	-0.13	-0.22	-0.84	-0.52	1.00	1.00

B

Annex 2 : Example output of parameterised wind model

In [29]:

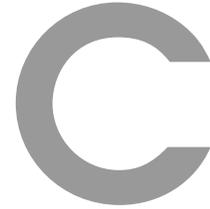
```
wind_manf_dict, wind_assembly_dict = wind_dict_creator(2000, print_details=True)
```

```
Nominal power : 2000 kW
Rotor diameter : 91.32639929505402 m
Hub height : 76.71205930786513 m
Tower weight : 203925.473526923 kg
Foundation weight : 1356413.7157780086 kg
Foundation reinforced steel: 27000.0 kg
Foundation concrete volume: 604.2789617172766 m3
Nacelle weight: 70807.84015999998 kg
Rotor weight: 48674.257255245786
Electronics: 3946.0
Total weight: 1683767.2867201772 kg
Input
  Nacelle
70807.84015999998
  Yaw - Drive
    Chromium steel M = 70807.84015999998 pctg = 0.01
4808379568385098 1048.5493735068214
'market for steel, chromium steel 18/8, hot rolled' (kilogram, GLO, None)
  Brake
    Chromium steel M = 70807.84015999998 pctg = 0.00
7404189784192549 524.2746867534107
'market for steel, chromium steel 18/8, hot rolled' (kilogram, GLO, None)
  Generator
    Cast iron M = 70807.84015999998 pctg = 0.04
087112760874287 2893.996270878827
'market for cast iron' (kilogram, GLO, None)
    Copper M = 70807.84015999998 pctg = 0.0119454261
85163978 845.8298279621692
'market for copper' (kilogram, GLO, None)
    Chromium steel M = 70807.84015999998 pctg = 0.10
726202934033606 7594.992628767744
'market for steel, chromium steel 18/8, hot rolled' (kilogram, GLO, None)
    Aluminium 0% recycled M = 70807.84015999998 pct
g = 0.010217781902185717 723.4990677197068
'market for aluminium, wrought alloy' (kilogram, GLO, None)
  Transmission
    Cast iron M = 70807.84015999998 pctg = 0.10
859478350149071 7689.362072383357
'market for cast iron' (kilogram, GLO, None)
    Rubber M = 70807.84015999998 pctg = 0.0049361265
22795033 349.5164578356072
'market for synthetic rubber' (kilogram, GLO, None)
    Chromium steel M = 70807.84015999998 pctg = 0.10
859478350149071 7689.362072383357
'market for steel, chromium steel 18/8, hot rolled' (kilogram, GLO, None)
  Hydraulic
    Chromium steel M = 70807.84015999998 pctg = 0.02
4680632613975163 1747.5822891780358
'market for steel, chromium steel 18/8, hot rolled' (kilogram, GLO, None)
    Lubricating oil M = 70807.84015999998 pctg = 0.00
2902442395403479 205.51567720733698
'market for lubricating oil' (kilogram, RoW, None)
  Yaw - ball bearings
    Low-alloy steel M = 70807.84015999998 pctg = 0.02
887634015835094 2044.6712783383018
'market for steel, low-alloyed, hot rolled' (kilogram, GLO, None)
  Axis
```

3019922206646 10835.010192903821 Low-alloy steel M = 70807.84015999998 pctg = 0.15
 'market for steel, low-alloyed, hot rolled' (kilogram, GLO, None)
 Bearing
 Cast iron M = 70807.84015999998 pctg = 0.01
 2389677572215532 877.2863091673739
 'market for cast iron' (kilogram, GLO, None)
 Chromium steel M = 70807.84015999998 pctg = 0.01
 2389677572215532 877.2863091673739
 'market for steel, chromium steel 18/8, hot rolled' (kilogram, GLO, None)
 Yaw - Brake
 Chromium steel M = 70807.84015999998 pctg = 0.00
 9872253045590066 699.0329156712144
 'market for steel, chromium steel 18/8, hot rolled' (kilogram, GLO, None)
 Frame
 Chromium steel M = 70807.84015999998 pctg = 0.27
 898987106837525 19754.670196868516
 'market for steel, chromium steel 18/8, hot rolled' (kilogram, GLO, None)
 Paneling
 Fiberglass M = 70807.84015999998 pctg = 0.06
 224455545244536 4407.402533307006
 'market for glass fibre reinforced plastic, polyamide, injection moulded' (k
 ilogram, GLO, None)
 Power supply
 620
 Cable
 Copper M = 620 pctg = 0.5387339530765826
 334.01505090748117
 'market for copper' (kilogram, GLO, None)
 HDPE granules M = 620 pctg = 0.2629482071
 7131473 163.02788844621514
 'market for polyethylene, high density, granulate' (kilogram, GLO, None)
 PP granules M = 620 pctg = 0.0088534749
 88933156 5.489154493138557
 'market for polypropylene, granulate' (kilogram, GLO, None)
 PVC impact resistant M = 620 pctg = 0.18
 946436476316955 117.46790615316512
 'market for steel, low-alloyed, hot rolled' (kilogram, GLO, None)
 Foundation
 Electronics
 3946.0
 Transformer
 Tin M = 3946.0 pctg = 0.004998000799680128
 19.722111155537785
 'market for tin' (kilogram, GLO, None)
 PVC impact resistant M = 3946.0 pctg = 0.05
 997600959616154 236.66533386645344
 'market for steel, low-alloyed, hot rolled' (kilogram, GLO, None)
 Aluminium 0% recycled M = 3946.0 pctg = 0.00
 03998400639744103 1.577768892443023
 'market for aluminium, wrought alloy' (kilogram, GLO, None)
 Copper M = 3946.0 pctg = 0.02998800479808077
 118.33266693322672
 'market for copper' (kilogram, GLO, None)
 HDPE granules M = 3946.0 pctg = 0.2698920431
 827269 1064.9940023990405
 'market for polyethylene, high density, granulate' (kilogram, GLO, None)
 Low-alloy steel M = 3946.0 pctg = 0.6297481007
 596961 2484.986005597761
 'market for steel, low-alloyed, hot rolled' (kilogram, GLO, None)
 Lead M = 3946.0 pctg = 0.004998000799680128

Annexe B. Annex 2 : Example output of parameterised wind model

19.722111155537785
'market for lead' (kilogram, GLO, None)
Tower
203925.473526923
Material
Low-alloy steel M = 203925.473526923 pctg = 0.99
48394495412844 202873.10583096984
'market for steel, low-alloyed, hot rolled' (kilogram, GLO, None)
Epoxy resin M = 203925.473526923 pctg = 0.00
5160550458715597 1052.3676959531576
'market for epoxy resin, liquid' (kilogram, RoW, None)
Rotor
48674.257255245786
Extender
Chromium steel M = 48674.257255245786 pctg = 0.21
08843537414966 10264.63928511986
'market for steel, chromium steel 18/8, hot rolled' (kilogram, GLO, None)
Blades
Fiberglass M = 48674.257255245786 pctg = 0.57
14285714285714 27813.861288711876
'market for glass fibre reinforced plastic, polyamide, injection moulded' (kilogram, GLO, None)
Hub
Cast iron M = 48674.257255245786 pctg = 0.21
768707482993196 10595.756681414048
'market for cast iron' (kilogram, GLO, None)
Onshore foundation
M_fond = 1356413.7157780086
concrete = 604.2789617172766m3
reinforced steel = 27000.0kg
'market for concrete, sole plate and foundation' (cubic meter, CH, None)
'market for reinforcing steel' (kilogram, GLO, None)
'market for road' (meter-year, GLO, None)
Land use
Conversion of meadows and pastures [m2] 1121.0
'conversion of meadows and pastures LSED wind' (square meter, IT, None)
Conversion to industrial area [m2] 121.0
'conversion to industrial area LSED wind' (square meter, IT, None)
Use of industrial area 4840.0
'use industrial area LSED wind' (square meter-year, IT, None)
Use of traffic area 40000.0
'traffic area use LSED wind' (square meter-year, IT, None)
Use of urban area 1000.0
'urban area use LSED wind' (square meter-year, IT, None)
Assembly not in kg
Cable
Copper 188.75764890434598
'market for copper' (kilogram, GLO, None)
HDPE granules 92.12986314641044
'market for polyethylene, high density, granulate' (kilogram, GLO, None)
PP granules 3.1020155941552336
'market for polypropylene, granulate' (kilogram, GLO, None)
PVC impact resistant 66.38313371492201
'market for steel, low-alloyed, hot rolled' (kilogram, GLO, None)
Assembly in kg
'market for wire drawing, copper' (kilogram, GLO, None) 1486
'market for explosive, tovox' (kilogram, GLO, None) 10
'market for sheet rolling, steel' (kilogram, GLO, None) 240714
'market for sheet rolling, aluminium' (kilogram, GLO, None) 725
'market for sheet rolling, chromium steel' (kilogram, GLO, None) 50200



Annex 3 : Supplementary figures from chapter 5

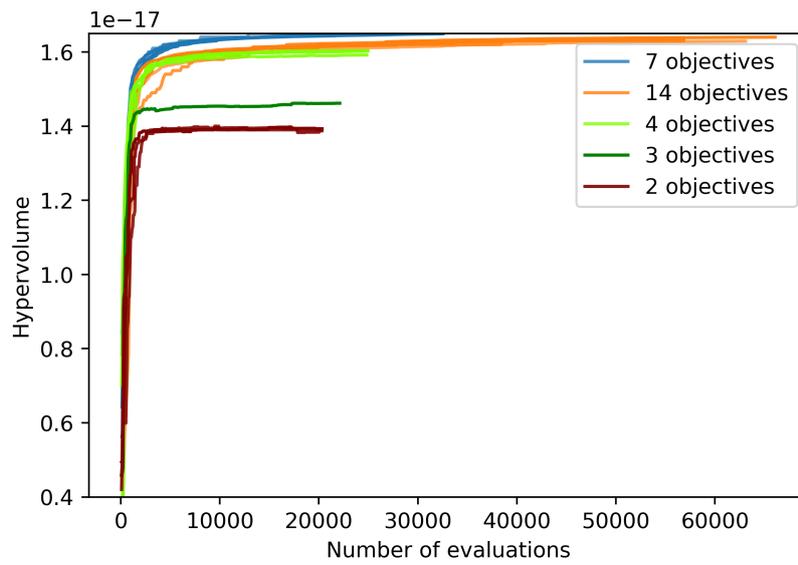


Figure C.1: Hypervolume indicator for different number of objectives with NSGA-III with extended number of evaluations

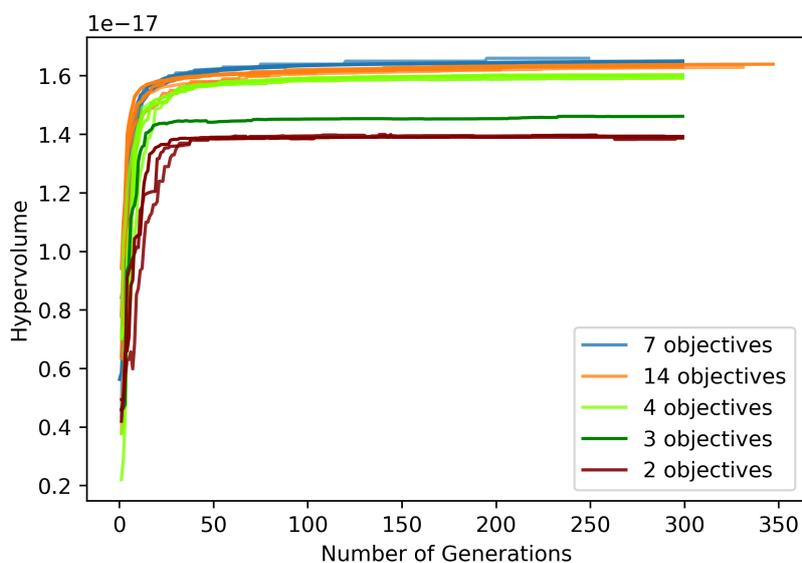


Figure C.2: Hypervolume indicator for different number of objectives with NSGA-III with number of generations

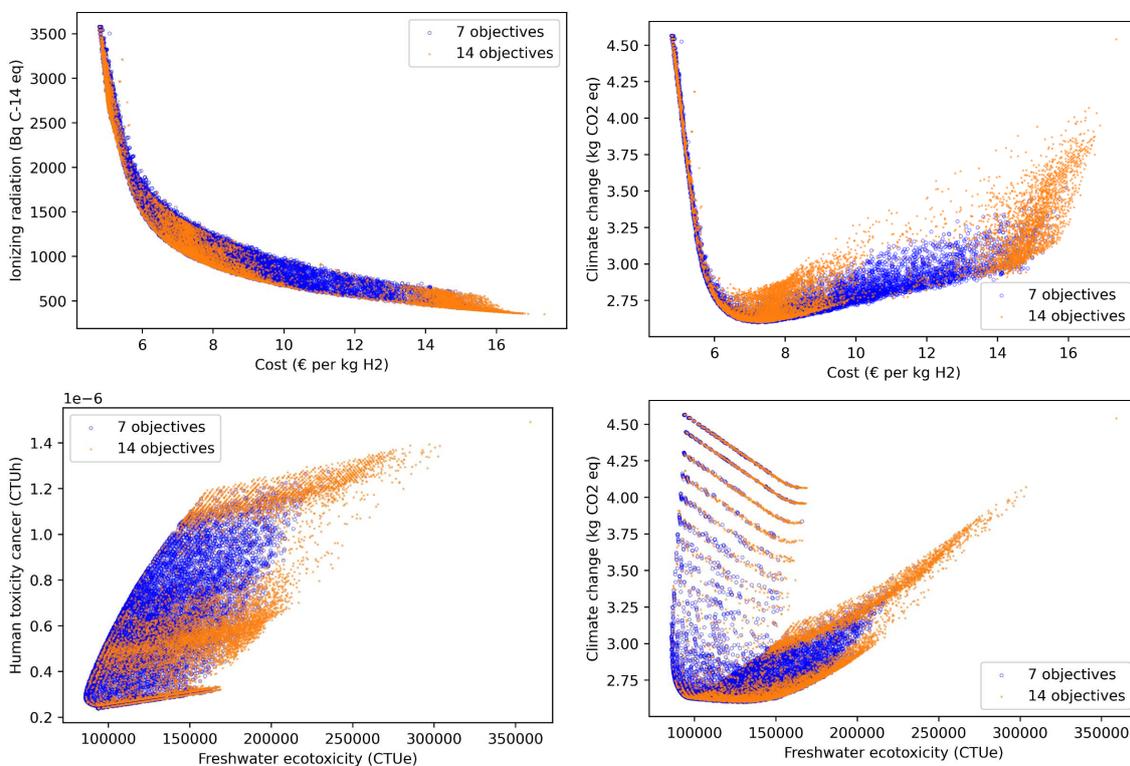


Figure C.3: Distribution of the pareto front obtained using 7 and 14 objectives shown in four figures of 2-dimensional plots

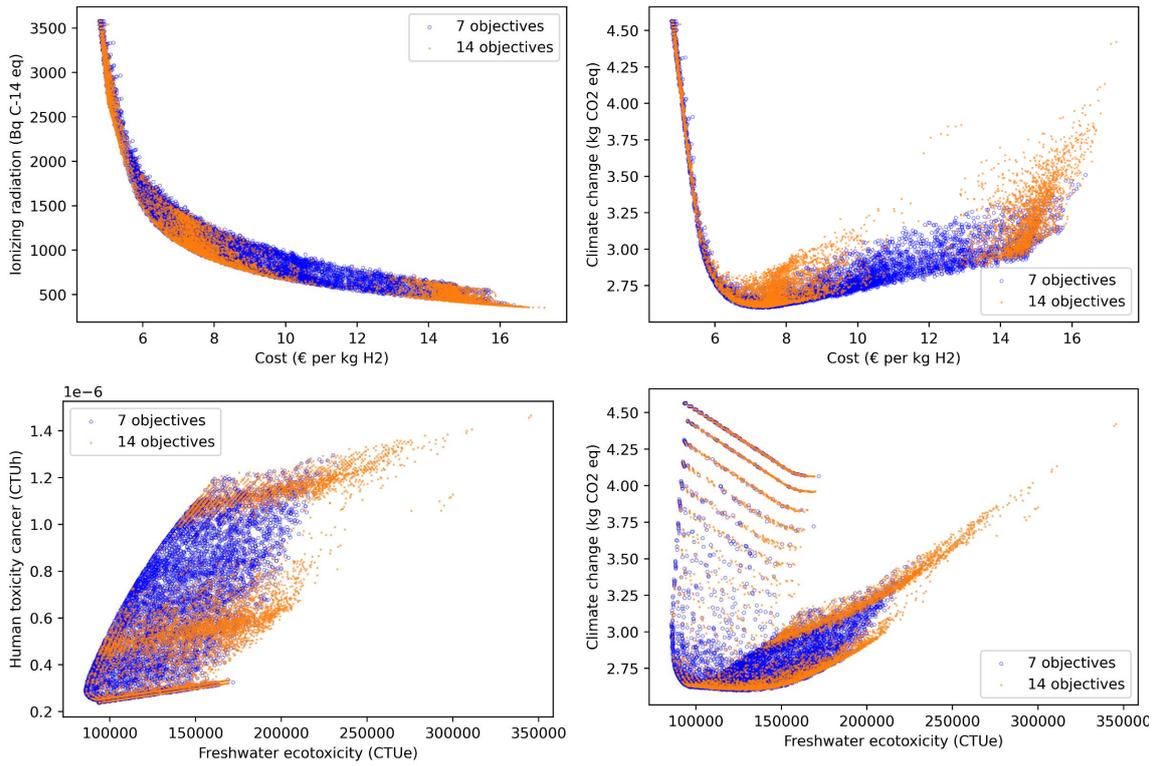


Figure C.4: Distribution of the pareto front obtained using 7 and 14 objectives shown in four figures of 2-dimensional plots

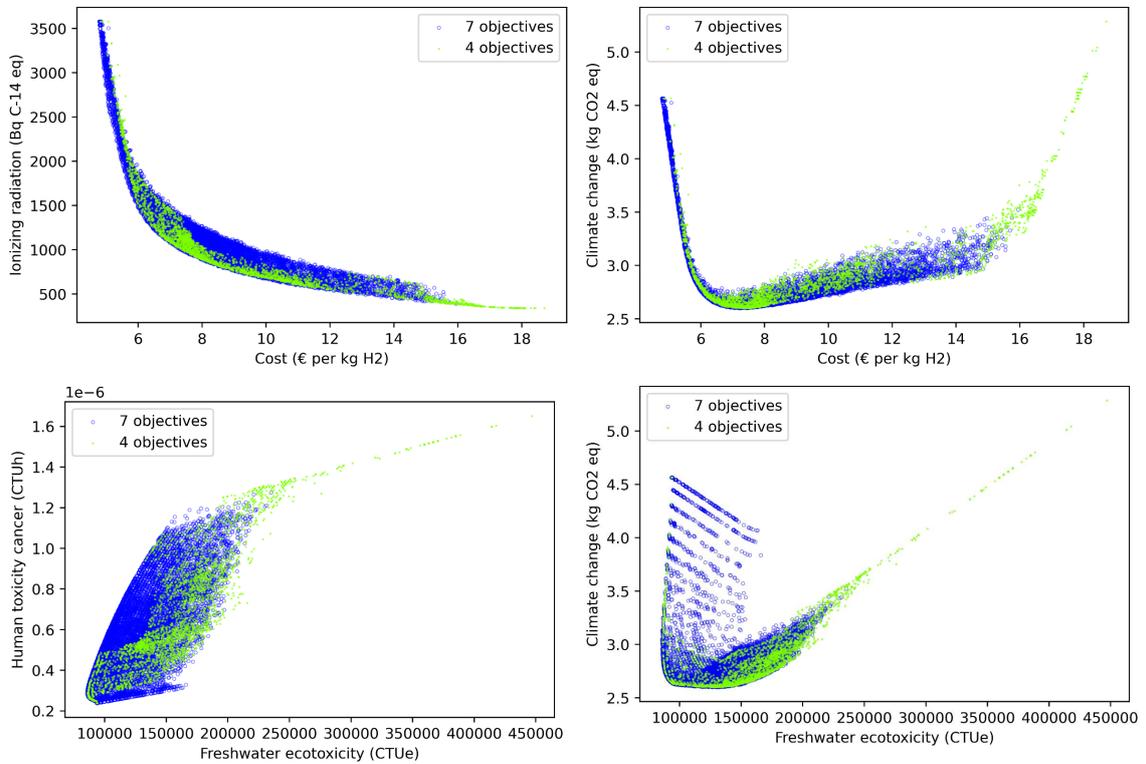


Figure C.5: Distribution of the pareto front obtained using 7 and 4 objectives shown in four figures of 2-dimensional plots

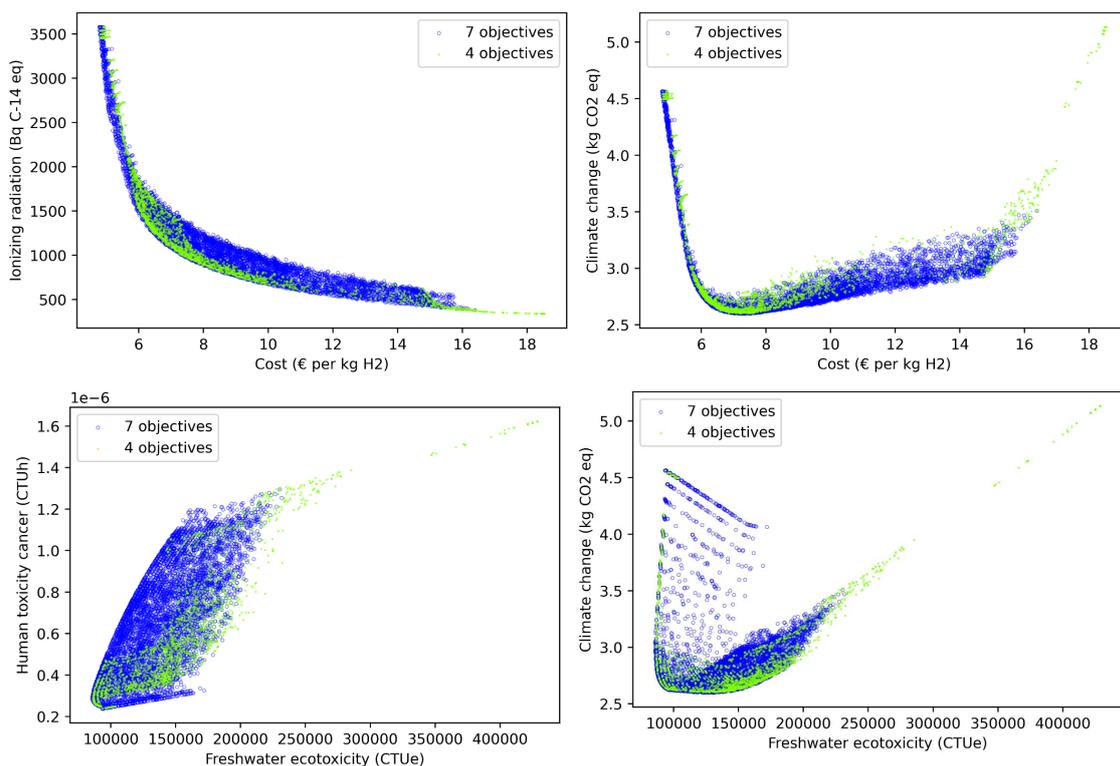


Figure C.6: Distribution of the Pareto front obtained using 7 and 4 objectives shown in four figures of 2-dimensional plots

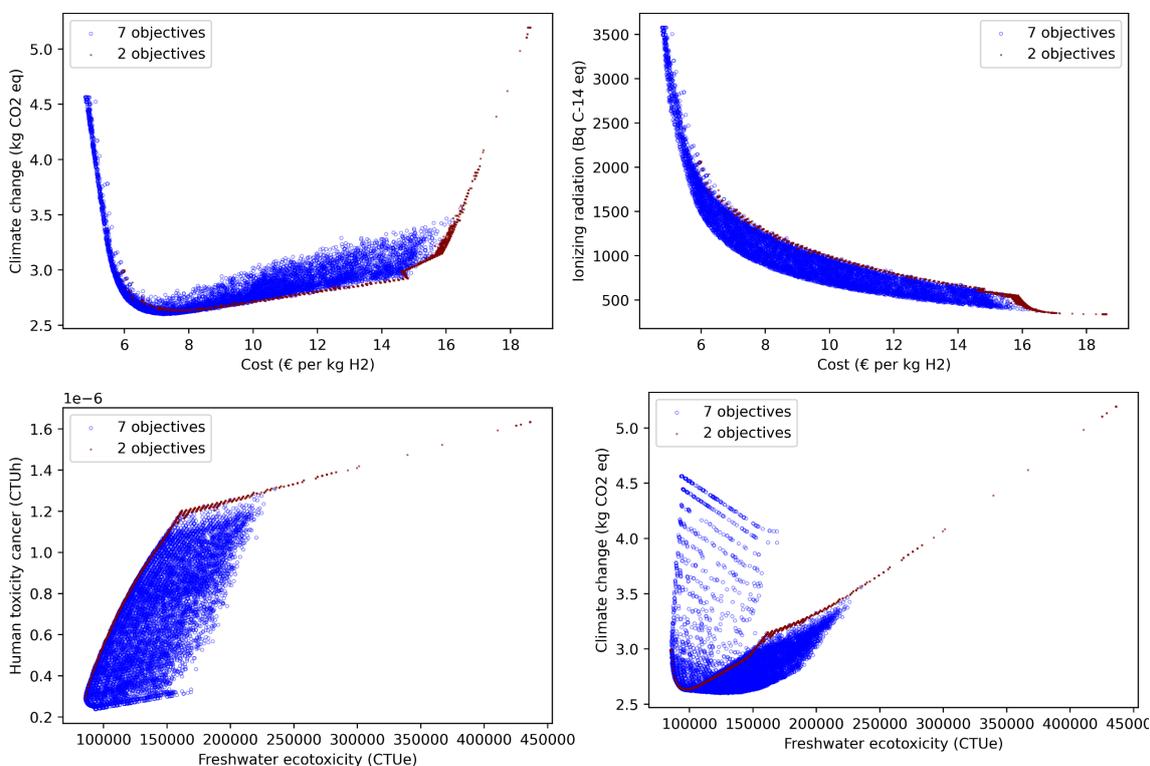


Figure C.7: Distribution of the Pareto front obtained using 7 and 2 objectives shown in four figures of 2-dimensional plots

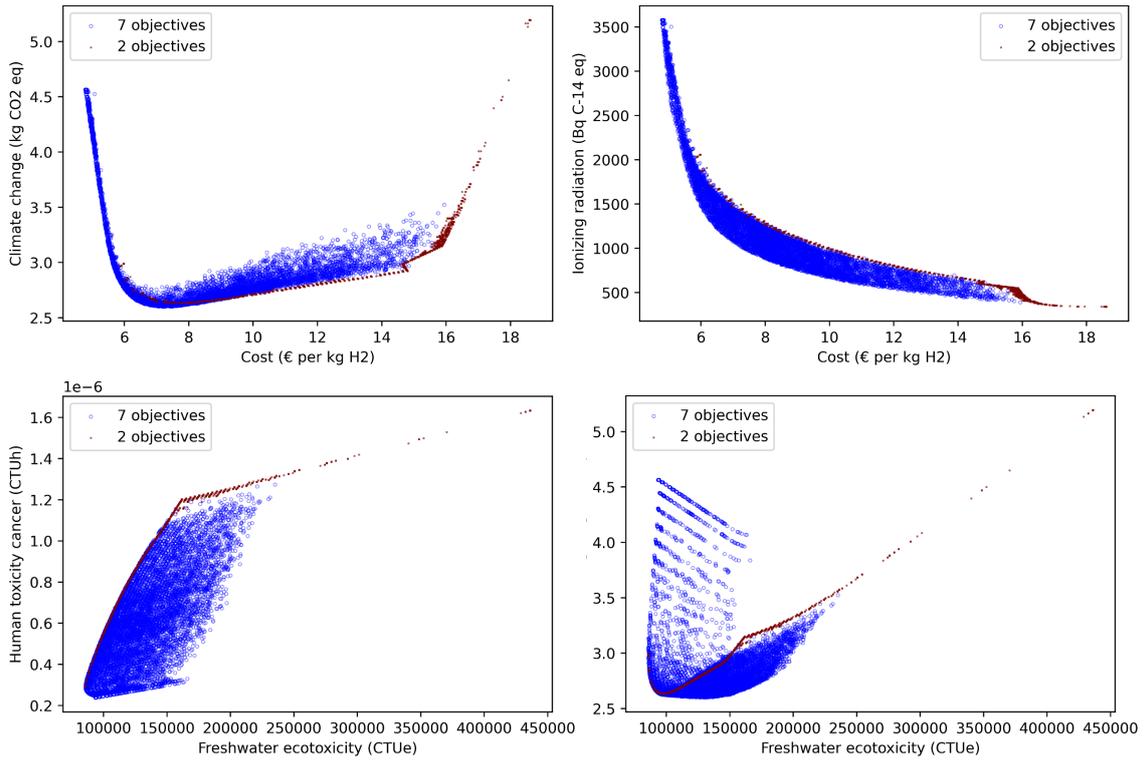


Figure C.8: Distribution of the pareto front obtained using 7 and 2 objectives shown in four figures of 2-dimensional plots

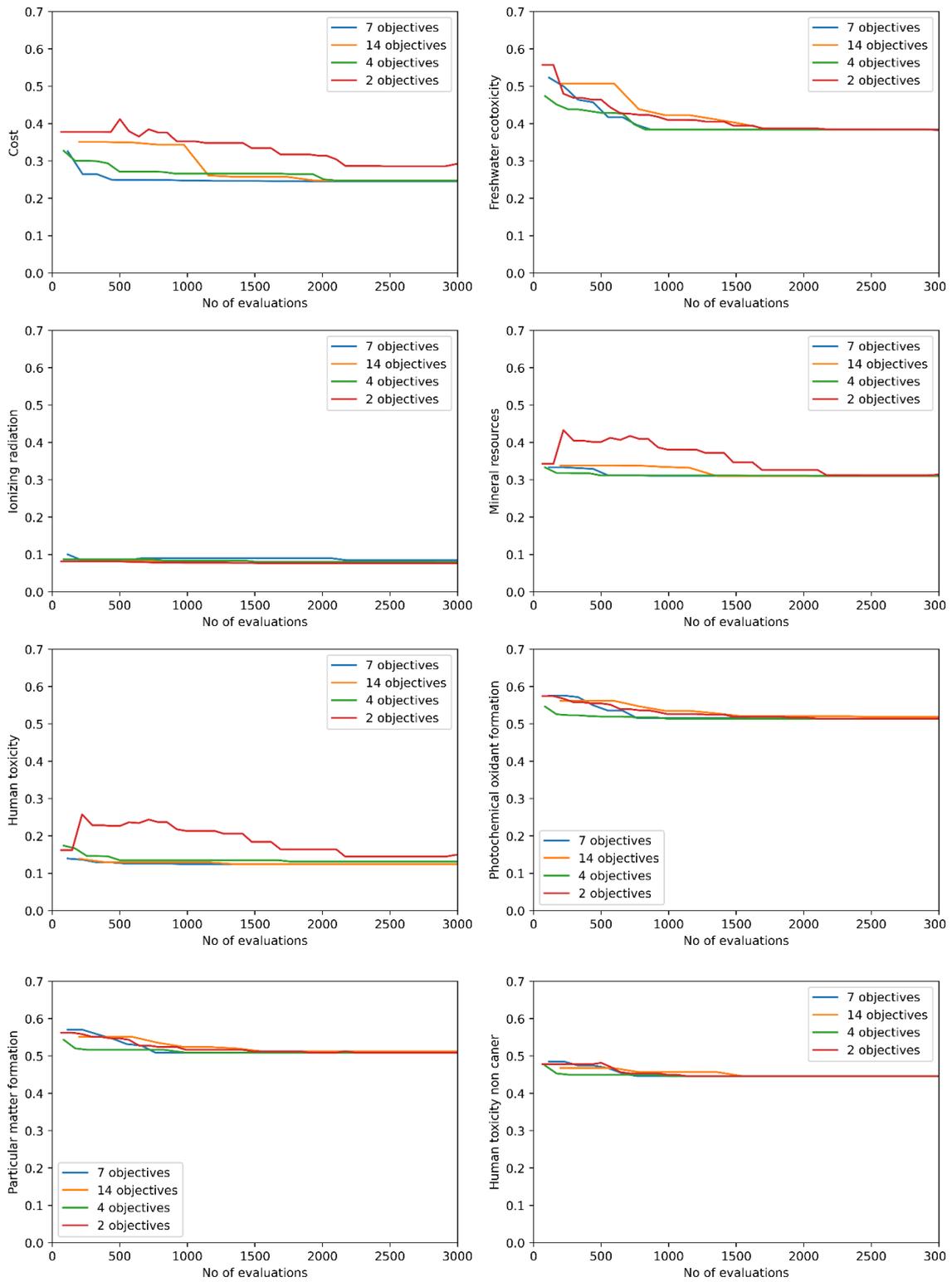


Figure C.9: Evaluations vs. standardised minimum point of indicators in the pareto front

Integration of environmental analysis in the assessments of hybrid energy systems

Résumé

Les technologies énergétiques émergentes sont souvent évaluées à une échelle pilote, ce qui permet d'étudier leurs performances dans des conditions réelles. On les appelle souvent systèmes énergétiques hybrides et ils se caractérisent par une offre/demande d'énergie variable, des vecteurs énergétiques multiples, etc. Les logiciels existants pour analyser ces systèmes ne permettent qu'une analyse environnementale simpliste qui pourrait être insuffisante pour évaluer les nouveaux systèmes énergétiques. Cette thèse vise donc à fournir une méthodologie générale pour inclure des indicateurs environnementaux dans la conception, la planification et l'évaluation des systèmes énergétiques hybrides. Dans ce contexte, trois contributions à la littérature sont proposées. Premièrement, pour faciliter la modélisation technico-économique et environnementale intégrée, un cadre pour développer des modèles paramétrés d'analyse du cycle de vie (ACV) des systèmes énergétiques hybrides est proposé. Deuxièmement, une approche d'optimisation pour la recherche de solutions à plus haute efficacité environnementale combinant la NSGA-III et la réduction objective est proposée. Troisièmement, une nouvelle approche de réduction d'objectif trouvée dans cette thèse est proposée, qui peut être appliquée à des cas simples.

Mots-clés : ACV; technologies émergentes; réduction des objectifs; optimisation multi-objectifs; systèmes énergétiques hybrides; analyse de sensibilité globale

Abstract

Emerging energy technologies are often assessed at a pilot scale, which allows the investigation of their performance in real world conditions. These are often called hybrid energy systems and are characterised by variable energy supply/demand, multiple energy vectors, etc. Existing software to analyse these systems only allow simplistic environmental analysis which might be insufficient to assess the new energy systems. This thesis thus aims to provide a general methodology to include environmental indicators in the design, planning and assessments of hybrid energy systems. Within this context, three contributions to the literature are proposed. First, to facilitate integrated techno-economic-environmental modelling, a framework to develop parameterised life cycle assessment (LCA) models of hybrid energy systems is proposed. Second, an optimisation approach to search solutions with higher environmental efficiency combining NSGA-III and objective reduction is proposed. Third, a new objective reduction approach found in this thesis is proposed, which can be applied to simple cases.

Keywords : LCA; emerging technology; objective reduction; multi-objective optimisation; hybrid energy systems; Global sensitivity analysis

