Science of the Total Environment 916 (2024) 169599

Contents lists available at ScienceDirect

Science of the Total Environment

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

Multiple criteria decision analysis to support the design of safe and sustainable chemicals and materials

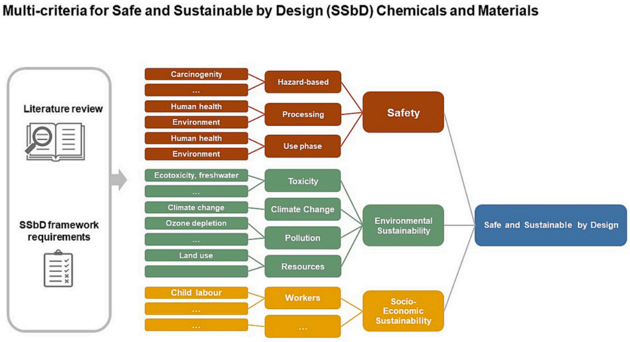
Luis C. Dias[a], Carla Caldeira[b,1], Serenella Sala[b,*]

[a] University of Coimbra, CeBER, Faculty of Economics, Av. Dias da Silva, 165, 3004-512 Coimbra, Portugal
[b] European Commission, Joint Research Centre, Via Enrico Fermi 2749, Ispra, Italy

HIGHLIGHTS

- Studies using MCDA in the context of safe and sustainable chemicals were revised.
- Characteristics of the safe and sustainable by design chemicals evaluation were identified.
- Different aggregation methods and approaches to consider uncertainty were explored.
- Multiattribute aggregation should be complemented with a detailed dashboard.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Jacopo Bacenetti

Keywords:
Sustainability
Life cycle assessment
Safety
Decision support
Chemicals
Materials
Multicriteria assessment

ABSTRACT

The development of safe and sustainable chemicals and materials is essential to achieve the Zero-Pollution Ambition for a Toxic Free Environment stated in the EU Green Deal. For that, criteria need to be defined and considered since early stage of development. A Safe and Sustainable by Design (SSbD) framework is proposed in an EU Recommendation suggesting the assessment of multiple safety and sustainability aspects of chemicals and materials leaving open how the evaluation and selection of the preferable option should be done. This paper presents a proposal with different options for the use of multiattribute aggregation in an evaluation procedure for the SSbD assessment of chemicals and materials. This proposal is based on i) a review of the literature focusing on Multi-Criteria Decision Analysis (MCDA) application in the SSbD context (i.e. applications considering simultaneously safety and sustainability attributes) and ii) the definition of requisites for MCDA to be applied to the SSBD framework. In the latter, an absolute rather than a relative assessment is preferred as it should be possible for an organization developing a new chemical or material to assess if it is SSbD, without needing to obtain data on all of its possible competitors. Moreover, rank-reversals caused by the introduction of other options are avoided, i.e., assessments of one alternative that depends on other alternatives being assessed simultaneously are not the most adequate. Different options for the aggregation of attributes at different levels are discussed as well as for the consideration of data quality in the evaluation procedure. Regardless the approach selected, the use of multiattribute aggregation does not rule out a richer dashboard presenting not only the overall aggregate result,

---

* Corresponding author.
E-mail address: serenella.sala@ec.europa.eu (S. Sala).
[1] Current address: Syensqo, Laboratory Research & Process - Sustainable & Innovative Processes, 85 rue des Frères Perret, 69192 Saint-Fons, France.

https://doi.org/10.1016/j.scitotenv.2023.169599
Received 29 July 2023; Received in revised form 21 November 2023; Accepted 20 December 2023
Available online 25 December 2023
0048-9697/© 2024 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

but also the results obtained in other levels of the hierarchy. Such complementary information is important to understand the strengths and weaknesses that an aggregate result might hide.

1. Introduction

One of the European Green Deal priorities is a Zero-Pollution Ambition for a Toxic Free Environment (EC, 2019). To achieve this objective, several initiatives have been developed, including the publication of a strategy on the sustainable use of chemicals: the Chemicals Strategy for Sustainability (CSS) (EC, 2020). A key action in the CSS is to develop EU 'safe and sustainable-by-design' (SSbD) criteria for chemicals, which should cover the different dimensions of sustainability, integrating safety, circularity and functionality of substances, advanced materials, products and processes throughout their lifecycle and by minimizing their environmental footprint (EC, 2020).

To support the development of SSbD criteria, the European Commission Joint Research Centre developed a framework (Caldeira et al., 2022a) that underpins the Commission recommendation on SSbD recently published (EC, 2022). The recommendation proposes the assessment of multiple safety and sustainability aspects of chemicals and materials without providing how one should deal with the multidimensional nature of this assessment for which decision support is needed. The review of existing multi-dimensional assessment frameworks for SSbD (Caldeira et al., 2022b) identified conceptual and operational frameworks, being the latter more limited and for which a procedure to support decision making is not provided.

A well-established approach for decision support is Multi-criteria Decision Aiding (or Analysis) (MCDA), which recognizes that decision makers may assess the objects of the evaluation (chemicals, materials, processes, etc.), usually named "alternatives" in MCDA, on multiple incommensurable dimensions. MCDA allows the participants in a decision process to include all of their perspectives, it allows focusing on a single dimension at a time, and it allows seeing the final decision as a compromise among different objectives, since typically no alternative is the best one on all the dimensions considered (Bouyssou, 1993). MCDA has been extensively used in the context of environmental management, with several books and reviews dedicated to this topic (e.g., Ehrgott and Stewart, 2010; Huang et al., 2011; Linkov et al., 2020; Linkov and Moberg, 2012).

The MCDA methodology encompasses several stages: a) Structuring, in which the scope of the assessment is delimited, the set (often a hierarchy) of objectives and attributes is structured, and the set of alternatives is developed; b) attribute-wise assessment, using direct measurements or defining suitable indicators; and c) multiattribute aggregation, in which the attribute-wise assessments are aggregated in order to produce a recommendation. This stage entails choosing the MCDA aggregation method, eliciting its parameters (including those related to weighing), obtaining overall results, and possibly performing sensitivity and robustness analyses to appraise how the results depend on the parameter values.

In the context of SSbD, multiattribute aggregation is to be used in an evaluation procedure for the safety and sustainability of chemicals and materials. Several dozens of methods for muiltiattribute aggregation have been proposed in the literature, as well as several taxonomies (Cinelli et al., 2020; Dias et al., 2019; Roy, 1996). Against this multitude of options it is relevant to identify the most appropriate multiattribute aggregation methods to be used in the context of SSbD. Therefore, building on a review on existing approaches and their potential applicability, the objective of this paper is to analyse the use of MCDA in support of the development of 'safe and sustainable-by-design' criteria for chemicals. The novelty of the present study lies in the specific review of the current MCDA methods applied to SSbD and the definition of requisites for MCDA to be applied in the SSbD framework (presented in Section 2) as a basis to propose possible options for the use of multiattribute aggregation in an evaluation procedure for the safety and sustainability of chemicals and materials (Section 3).

2. Background: literature review and requisites for the framework

To build a proposal with possible options for the use of multiattribute aggregation in an evaluation procedure for the SSbD assessment of chemicals and materials, two initial steps were taken. Step one was a review of the literature focusing on MCDA application in the SSbD context (i.e. applications considering simultaneously safety and sustainability attributes), presented in Section 2.1. The second step was the definition of requisites for MCDA to be applied to the SSbD framework proposed by Caldeira et al. (2022a), which are presented in Section 2.2.

2.1. Review of literature on MCDA applied to safety and sustainability assessments

MCDA has been highlighted as a key instrument for sustainability assessment in general, as discussed in major works and reviews by Munda (2005), Cinelli et al. (2014), Ibáñez-Forés et al. (2014), Diaz-Balteiro et al. (2017), or Lindfors (2021). These works highlight some of the strengths of MCDA concerning sustainability assessments, which include its alignment with the multidimensionality intrinsic to the sustainability concept, its potential to engage and to dialogue with stakeholders, its transparency and its emphasis on finding compromise solutions.

These works also highlight the methodological diversity of the MCDA field, acknowledging the difficulty of suggesting a single MCDA method as the best one for all situations. Some criteria have been suggested to guide this choice, such as scientific soundness, and feasibility (Cinelli et al., 2014), among many other criteria that are being proposed in the context of projects, most recently in Orienting Huysveld et al. (2021) and GLAM (Cinelli et al., 2022) projects.

In a recent review, Lindfors (2021) covers 280 articles, showing the increasing popularity of MCDA in sustainability assessments. Most of the applications addressed the energy sector, the construction sector, and the agri-food sector. Closer to the SSbD focus, 13 studies in the chemical industry plus two in nanotechnology were reported. Lindfors presents the choices made in these studies concerning attributes, normalization, weighting, and aggregation method (as Ibáñez-Forés et al. (2014) and Diaz-Balteiro et al. (2017) did earlier). In many cases, these choices were not explained by the respective authors, motivating Lindfors to call for increased transparency about methodological choices in published studies. Another important issue mentioned by Lindfors and other authors is the use of tools to address uncertainty, both related to value judgments and to data quality.

Also in relation with sustainability assessment, MCDA is often used in combination with Life Cycle Assessment (LCA) (Dias et al., 2019; Zanghelini et al., 2018), being considered as a means of weighting impact indicators at the interpretation stage (Pizzol et al., 2017; Roesch et al., 2020). MCDA is mostly associated to panel weighting, often involving participatory MCDA approaches (e.g., (Marttunen and Suomalainen, 2005; Munda, 2004; Mustajoki et al., 2004), which is a common type of weighting together with other types such as distance to target weighting, monetary weighting, etc. (Lindfors, 2021; Pizzol et al., 2017; Roesch et al., 2020; Sala et al., 2018).

MCDA has been used to support sustainability assessment concerning many different kinds of systems, products and services. Other studies have focused on safety rather than sustainability, and some studies addressed both safety and sustainability. A search for MCDA, sustainability or safety, and chemicals, substances and materials in the Scopus database with the *Query: (multicriteria OR multi-criteria OR "multiple criteria" OR mcda OR mcdm OR multiattribute) AND (chemical OR material OR substance) AND (safe* or sustainab*) in Title, Abstract or Keywords* returned almost 1700 results (17/07/2023), with numbers increasing in recent years. Among these, 1243 results are articles published in journals. Focusing on journal articles, those that did not explicitly assess chemicals or materials were excluded. Excluded studies deal with topics such as supply chain management (e.g., supplier selection, transportation, location), waste management (recycling, materials recovery, remediation), farming, and building design or retrofitting, to name some of the most common purposes.

This process resulted in 110 papers closer to the SSbD focus to be further analysed (presented in Supplementary Information (SI), merely as examples of the variety of existing efforts). Most applications of MCDA for sustainability or safety assessment of materials and chemicals address materials for the construction sector (36 studies, addressing building materials, road pavements, etc.). The other main groups identified were a group on chemical industry processes (23 studies), often to produce well-known substances such as hydrogen, as well as a group on materials for the automotive and aviation sector (11 studies), and a group on fuels (8 studies).

Among the 110 reviewed studies, only 27 consider both safety and sustainability, use MCDA, and focus on chemicals or materials. Table 1 presents these studies, indicating their purpose, the chemicals or materials assessed, the life cycle stages encompassed, the indicators used in the assessment, normalization used, MCDA method, and weighting process. More than half of these studies assess chemical industry processes (13 studies) or (bio)fuels (five studies). With only one exception (van Dijk et al., 2022), these studies make comparative assessments, either comparing different chemicals or materials, or comparing different ways to produce a given chemical.

Only seven studies mention explicitly that they include several stages of the chemical or material's life cycle, either covering the entire life cycle or adopting a cradle to gate scope. Eleven studies focused on the production (manufacturing) stage, one study focused on the use stage, another one on the use and end of life. The remaining ones are not explicit about the life cycle stages considered. Three of the studies assessed only environmental and safety indicators; most of the studies included additional technical or economic indicators, the latter being more common. Concerning environmental indicators, those related with Global Warming Potential (GWP) were the most common. A few studies used comprehensive life cycle impact assessment methods, namely ReCiPe (Goedkoop et al., 2009; Huijbregts et al., 2016) and Eco-Indicator (Goedkoop et al., 2009), either presenting multiple indicators, or presenting a synthesis score (in the case of Eco-indicator). Concerning safety, the chosen indicators are very diverse, in some cases computing a single safety index or risk score, and in other cases presenting multiple context-specific safety indicators.

An article by Stoycheva et al. (2022), not included in Table 1 for not featuring an environmental assessment, is noteworthy for its focus on social life cycle assessment (S-LCA) to support SSbD decision making. Rather than proposing comparative assessments, it aims at informing self-assessments (screening) of chemicals under development, with a focus on nanomaterials.

Concerning the use of MCDA methods, two approaches stand out: Analytic Hierarchy Process (AHP) (Forman and Gass, 2001; Saaty, 2005) was used in eleven studies, sometimes only to obtain weights that are then inputs to other MCDA methods; Weighted averages, including Multi-attribute Utility Theory (MAUT) (Keeney, 2006; Keeney and Raiffa, 1993), were used in four studies. Other methods were used less frequently.

Four studies did not actually perform an MCDA aggregation, using charts (e.g., spider charts, bar charts) to compare side by side the scores or the rankings of the alternatives compared on multiple indicators. In the study of Tahmid and Syeda (2023), the absence of an MCDA aggregation was justified by the fact that one of the alternatives was the best in the two criteria used, and the authors suggest using additional criteria (namely costs) and TOPSIS to perform an aggregation. Other studies are not included in Table 1 for not mentioning MCDA, and therefore not being encompassed by the search keywords. A recent noteworthy example is the article by Pizzol et al. (2023) addressing early-stage screening of multi-component nanomaterials. These authors proposed and applied a framework to compare qualitatively whether a material is better than a benchmark (yes, no, not known) over multiple safety, environmental, economic and social items, allowing to display the proportion of items where the new product is superior. As most of the methods used require that indicators are given in commensurable scales, a normalization operation is performed, the most common one being a max-min ratio normalization, e.g., giving for each indicator a normalized score of 0 to the worst alternative being assessed and a normalized score of 1 to the best one. It should be highlighted that all these studies were comparative assessments comparing multiple alternatives. Indeed, methods such as AHP and PROMETHEE (Brans et al., 1986), as well as the use of internal normalizations based on the best and worst observed impacts, makes no sense when only one (or even two) alternatives are being assessed.

The weights used in the MCDA aggregation, when present, usually reflect the perspectives of experts (in seven of the studies) or the perspectives of the authors of the study (in seven of the studies), or using a default weighing vector (e.g., equal weights). In the study by Purker et al. (2023), multiple stakeholders were surveyed to obtain weights.

2.2. Requisites multiattribute aggregation in the SSbD framework

The SSbD framework proposed in Caldeira et al. (2022a) defines a preliminary set of requisites to be taken into account when suggesting an overall evaluation procedure for SSbD. Table 2 lists the main requisites and their implications.

A key aspect of the framework is the hierarchical approach in which the evaluation procedure is underpinned by a hierarchical principle and the chemical/material should be considered SSbD if passing the criteria defined for safety and environmental sustainability. Therefore, the aggregation approach should respect the predefined hierarchy and not allow trade-offs allowed between safety and environmental performance. One dimension cannot compensate for weaknesses on the other dimension. If the minimum criteria for safety dimension are not met, then the chemical/material cannot be considered as SSbD.

A criterion should be defined as an aspect with an assessment method and a minimum or maximum threshold or target values, on which a decision can be based. Ideally, each attribute is associated with thresholds to act as classification criteria. This is more straightforward for safety assessments in which threshold values are normally available, but not usually done in environmental LCA. To overcome this, the ambition of the SSbD is to move from relative (safer and more sustainable) to absolute (safe and sustainable) improvements, ensuring that chemicals and materials are produced and used without exceeding acceptable boundaries. The procedure should evaluate each chemical or material based on its own merits (absolute evaluation independent from other chemicals or materials being assessed). The outcome of this assessment can be reflected in a qualitative level (rating) or having numerical values to be compared with thresholds.

The result of the evaluation can be expressed either as a class of SSbD (poor, good, very good) or with a numerical score derived from the combination of the individual scores in each aspect. The result can be provided on an ordinal scale as a qualitative level, i.e. a rating (Colorni and Tsoukiàs, 2021). Alternatively, it can be a continuously varying numerical value.

The fact that the evaluation procedure should be applied to new chemicals and materials or to existing ones implies that the data availability and quality can vary widely among attributes, leading to the need of considering data quality and uncertainty when taking decisions regarding the development of the chemicals.

Table 1
Previous MCDA studies addressing sustainability and safety of chemicals or materials.

Authors (Year)	Purpose	Chemical/Material	Life cycle stages	Environmental assessment	Safety assessment	Other assessments	Normali-zation	MCDA method	Weighting
(Abdel-Basset et al., 2021)	Comparison of hydrogen production options	Hydrogen	Production	Contaminants emission, land requirements, rigid waste generation (Qualitative assessments from experts)	Influence in public health (Qualitative assessments from experts)	Economic, technical, and social (Qualitative assessments from experts)	N/A	Analytic Hierarchy Process (AHP)-COPRAS-EDAS	Weights provided by experts
(Acar et al., 2018)	Comparison of alternative hydrogen production options	Hydrogen	Production	GHG emissions, land use, water discharge quality, and solid waste (Qualitative assessments from experts)	Impact on public health (Qualitative assessments from experts)	Economic, technical, availability and other social (Qualitative assessments from experts)	N/A	hesitant fuzzy AHP	Weights provided by three experts from hydrogen production industry
(Acar et al., 2022)	Comparison of fuel cells	polymer electrolyte membrane, alkaline, phosphoric acid, molten carbonate, and solid oxide fuel cells	Not explicit	GHG, Land use, water discharge quality, solid waste generation (Qualitative and comparative assessments from three experts)	Human Health Impact (Qualitative and comparative assessments from three experts)	Economic, technical, and social	N/A	AHP (fuzzy variant)	Weights provided by experts
(Alkhatib et al., 2021)	Comparison of Hybrid Solvents for acid gas removal	mixtures of chemical solvents	Cradle to grate	LCA indicators (Cumulative energy demand, GWP, Eco-indicator 99) - assessment method FineChem Tool	Health and safety hazards	Relevant physicochemical solvent properties - using Polar soft-SAFT	N/A	Single-criterion rankings are compared, without an overall aggregation	N/A
(Banimostafa et al., 2012)	Comparison of chemical routes during early process design	4-(2-methoxyethyl)-phenol (MEP) and methyl methacrylate (MMA)	Not explicit	CED (Ecoinvent)	Risks (toxicity, safety, etc)	–	Internal (max-min normalization and other)	principal component analysis (PCA) based	Driven by the PCA
(Crivellari et al., 2021)	Comparison of alternative processes for synthetic methanol synthesis	Methanol	Production	Technological performance (energy efficiency) Env. Performance (Levelized GHG Emissions- the averaged emission over the lifetime of the process scheme)	Safety (Inherent hazard of the production unit with respect to humans)	Economic performance	Between zero (undesired) and 1 (desired) comparing the actual indicator with respect to a given target value	AHP	Four perspectives: Individualist, Egalitarian, Hierarchist, and Equal weighting
(Dinh et al., 2009)	Comparison of biodiesel production alternatives	Biodiesel from 5 different feedstocks	All	LCIA (GHG, water, land use),	methanol ratio, flash point	Economic and technical	N/A	AHP	Weights defined by the authors
(Iranfar et al., 2023)	Comparison of construction materials	Eight traditional and new construction materials	Production	Water, Energy, Recyclability, Sustainability (qualitative assessments)	Safety for workers (qualitative assessment)	Resource availability and technical	Internal (division by vector norm)	AHP (fuzzy), TOPSIS, VIKOR, WASPAS	Weights defined by the authors
(Janošovský et al., 2022a, 2022b)	Comparison of hydrogen production processes	Hydrogen	Not explicit	Environmental Impact (C factor, Eco-Indicator 99)	Process Safety (Process Route Index, Comprehensive Inherent Safety Index)	Economic, Material and energy utilization	Internal (max-min normalization)	AHP	Weights defined by the authors
(Jia et al., 2016)	Comparison of chemical processes	chemicals in general, example for ethanol	Production	Environmental (global warming potential (GWP), photochemical oxidation potential (PCOP), ozone depletion potential (ODP), acidification potential (AP), eutrophication potential (EP), human toxicity potential by ingestion	Safety (inherent safety, process safety)	Economic	Integer 1–5 scores	Analytic Hierarchy Process (AHP)	Weighs provided by decision makers and domain experts

(*continued on next page*)

Table 1 (*continued*)

Authors (Year)	Purpose	Chemical/Material	Life cycle stages	Environmental assessment	Safety assessment	Other assessments	Normali-zation	MCDA method	Weighting
				(HTPI), human toxicity potential by exposure both dermal and inhalation (HTPE), aquatic toxicity potential (ATP), and terrestrial toxicity potential (TTP).					
(Li et al., 2011)	Comparison of biodiesel production processes in early design stages	Biodiesel produced using two processes	Production	Waste Reduction (WAR) Algorithm	Enhanced Inherent Safety Index	Economic, efficiency	Internal (max-min normalization)	Weighted average	Equal weights
(Liew et al., 2013)	Comparison of biodiesel production pathways during R&D	8 Biodiesel pathways using different feedstock and transesterification processes	Not explicit	Inherent Environmental Toxicity Hazard	Prototype Index of Inherent Safety, Inherent Occupational Health Index	–	Internal (max-min normalization)	Maxmin (fuzzy variant)	N/A
(Limleamthong et al., 2016)	Comparison of molecules to screen inefficient chemicals and provide improvement targets	Chemicals, example for amine-based solvents for CO2 capture	Cradle to Gate for LCA impacts	Eco-Indicator 99	vapor pressure, fire & explosion mobility, acute toxicity	Other properties	N/A	Data Envelopment Analysis	Weights computed from the data (a different vector for each alternative)
(Morales et al., 2016)	Assess alternative ways of producing Succinic acid	Succinic acid	Cradle to gate	Env - LCA (non-renewable cumulative energy demand (CED), the global warming potential 100a (GWP) and eco-indicator 99 (EI-99))	Env, Health, Safety Hazard assessment (physical and chemical properties to describe eleven dangerous properties; assess the damage to the environment, the workplace and surroundings in accidental scenarios and damage to workers' health as a result of long-term exposure to the chemicals in the process)	Eco - Operating costs	N/A	Single-criterion rankings are compared, without an overall aggregation	N/A
(Moretti et al., 2017)	Comparison of cement powders	Portland-pozzolana cement	Production	Global Warming Potential, Ozone Depletion Potential, Photochemical Ozone Creation Potential, Eutrophication Potential, Non-hazardous Waste, Hazardous Waste, Renewable resources with energy content, Non-renewable resources with energy content, Electricity, Water consumption	Risks for construction industry workers: noise, whole body vibrations (WBV), hand-arm vibrations (HAV), exposure to allergizing substances, exposure to Cr(VI) (hexavalent chromium), and exposure to free crystalline silica (FCS)	socio-economic	N/A	Analytic Hierarchy Process (AHP)	Weights provided by technicians from different backgrounds experts in the fields of environment, human health and economy
(Narayanan et al., 2007)	Comparison of biodiesel production alternatives (incl.	Biodiesel alternatives (4 raw materials, 3 catalysts, and other process variables)	All	LCIA (GWP, Ozone, Acid., Ecotoxicity etc)	qualitative risk assessment matrix (personnel, environment)	Economic and technical	N/A	Analytic Hierarchy Process (AHP)	Weights defined by the authors

(*continued on next page*)

Table 1 (*continued*)

Authors (Year)	Purpose	Chemical/Material	Life cycle stages	Environmental assessment	Safety assessment	Other assessments	Normali-zation	MCDA method	Weighting
	Raw materials and process)								
(Posada et al., 2013)	To assess bioethanol as a chemical building block for biorefineries vs. fossil fuel	12 bioethanol-based products (Ethylene, Propylene, etc.)	All	LCIA (CED, GHG)	Environmental-Health-Safety index (EHSI) and risk aspects	Economic	Internal (based on the worst impacts)	Weighted average	Weights provided by experts
(Purker et al., 2023)	To assess nano-materials	Two studies: Carbon nanofibers, nano-materials in tire rubber	All	Envirnmental LCA (details not provided)	Worker safety (details not provided)	Economic, social and technical aspects	External considering a “neutral” option	ELECTRE TRI / SMAA	Survey to stakeholders
(Saavalainen et al., 2015)	Comparison of Chemical Processes	dimethyl carbonate	Production	feedstock renewability, energy intensity, waste generation, CO2 balance	chemical safety indicator	economic indicators, process conditions and innovation potential	Internal (max-min normalization)	Only visualization (spidergram)	N/A
(Samani et al., 2015)	Comparison of materials for buildings	polymer matrix composite sandwich materials	Production	Envir, LCA (ReCiPe)	Fire performance	mechanical, thermal and acoustic properties, costs	ReCiPe	PROMETHEE II	ReCiPe and authors
(Seker and Aydin, 2022)	Comparison of processes to obtain hydrogen from H_2S	Hydrogen	Production	Ecological feasibility, Energy requirement (Qualitative assessments from experts)	Environmental and occupational safety (Qualitative assessments from experts)	Economic and technical aspects	Internal (max-min normalization)	SWARA (fuzzy) and IVIF-WASPAS	Group of experts
(Shi et al., 2014)	Comparison of snow and ice control chemicals for highways	13 rock salts,3 IceSlicer products, 8 salt brines, and a corrosion-inhibited magnesium chloride	Use	Lethal concentration (ALC) for aquatic species; chemical oxygen demand; biochemical oxygen demand; risk to air quality (particle matter)	Chloride anion emission	Economic and technical performance	Internal (max-min normalization)	Weighted average	Equal weights
(Simanovska and Grigale-Soročina, 2016)	Comparison of nail polish products	Nail polish, Gel polish, Hybrid polish	All, but no LCA	Use of fossil fuel	GHS classification, EDS classification	Durability	N/A	Only visualization (spidergram)	N/A
(Sun et al., 2023)	Compare distillation processes for complex azeotropic mixtures	Three processes to separate ACN/EtOH/water	Production	CO_2 emissions	Inherent safety (PRI index)	Costs	Internal (division by vector norm)	TOPSIS	Entropy weighting
(Tahmid and Syeda, 2023)	Propose a framework to compare chemical processes	Four solvent alternatives for palm oil recovery	Not explicit	Green degree value (a weighted sum)	Total inherent safety score	–	Internal (division by maximum) for green degree value	Single-criterion rankings are compared, without an overall aggregation	Used for green degree value, but not explicit
(van Dijk et al., 2022)	Workflow to facilitate the chemical redesign for reduced persistency	Chemicals, organophosphate chemical triisobutylphosphate (TiBP)	N/A (focus on use and end of life)	QSAR models from EPIsuite and VEJA platform for properties P (biodegrability), B (BCF), M (Koc)	QSAR models from EPIsuite and VEJA platform for property T (Non-Mutagenicity, EDC)	Technical	Internal (max-min normalization)	Multi-Attribute Utility Theory (MAUT)	Weights defined by the authors

Table 2
Requisites underpinning an overall evaluation procedure for SSbD (adapted from Caldeira et al. (2022a)).

Requisites	Implications
1. The evaluation procedure can be applied to new chemicals and materials or to existing ones.	By including new chemicals and materials, data quality can vary widely among attributes.
2. The evaluation procedure shall take into account the lack of data and data uncertainty.	Data quality needs to be assessed.
3. The result of the evaluation can be expressed either as a class of SSbD (poor, good, very good) or with a numerical score[a] derived from the combination of the individual scores in each aspect.	The result can be provided on an ordinal scale as a qualitative level, i.e. a rating (Colorni and Tsoukiàs, 2021). Alternatively, it can be a continuously varying numerical value.
4. A criterion is defined as an aspect with an assessment method and a minimum or maximum threshold or target values, on which a decision may be based.	Each attribute is associated with thresholds to act as classification criteria. A qualitative level (rating) is obtained for each attribute. The numerical values to be compared with the thresholds might be also available.
5. The ambition of the SSbD is to move from relative (safer and more sustainable) to absolute (safe and sustainable) improvements ensuring that chemicals and materials are produced and used without exceeding acceptable boundaries.	The procedure should evaluate each chemical or material based on its own merits (absolute evaluation independent from other chemicals or materials being assessed).
6. The chemical/material should be considered SSbD if passing the criteria defined for safety and environmental sustainability.	No trade-offs allowed between safety and environmental performance. One dimension cannot compensate for weaknesses on the other dimension.
7. The evaluation procedure is underpinned by a hierarchical principle (p.46). A 'step score' and an 'overall SSbD score' could be developed considering the combination of scores. If the minimum criteria for safety dimension are not met, then the chemical/material cannot be considered as SSbD.	The aggregation approach should respect the predefined hierarchy. A poor safety assessment cannot be overridden by the environmental assessment.

[a] The expression "numerical score" is somewhat ambiguous because it often refers to a 0–4 scale of integer values in which the numbers are not quantities (cardinal) but just labels (e.g., labels E, D, C, B, A could be used instead).

The literature reviewed in Section 2.1 indicates much progress is still needed. To start, only 27 MCDA studies addressed both safety and sustainability indicators for chemicals or materials, and most of these were comparative studies (i.e., assessing competing alternatives). Also, only seven studies encompassed the whole life cycle of the chemical or materials assessed. Finally, an actual MCDA aggregation was not performed in all of the studies. A variety of multiattribute aggregation methods was used in the remaining cases, warranting a closer examination of their adequacy for the framework's requisites, as discussed in the following section.

3. Possible options for multiattribute aggregation for decision support in the SSbD framework

Building on the literature review and on the requisites for MCDA applied to the SSbD framework, the main methodological options are herein discussed and options on how to aggregate the multiple aspects encompassed by this framework are presented. As illustrated in Fig. 1, the SSbD framework presents a hierarchy with different levels that can be aggregated. Level one corresponds to the sustainability dimensions

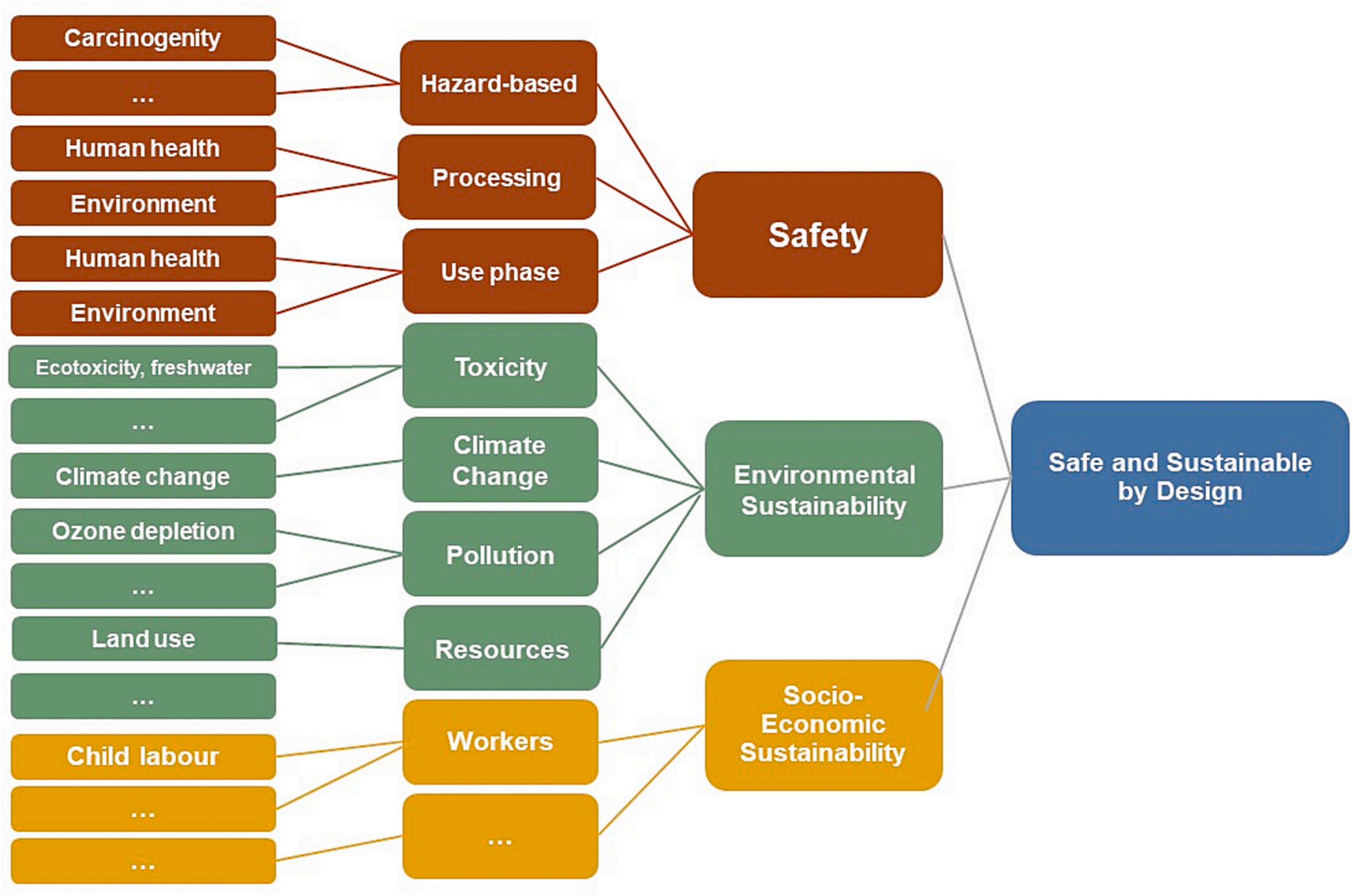


Fig. 1. Aggregation as a hierarchy based on the framework by Caldeira et al. (2022a).

considered in the framework (safety, environmental, and socio-economic), the second level are the steps within each dimension (e.g. Safety, step 1 assesses the hazard properties of the chemical/material under assessment). Finally, the third level refers to the specific aspects, for example under step 1 carcinogenicity is assessed.

Sections 3.1 and 3.2 discuss multiattribute aggregation methods that can be used at different levels of the hierarchy (Fig. 1), based mainly on the requisites and implications identified in Table 2. Considerations of data quality are addressed in Section 3.3.

3.1. MCDA considerations

3.1.1. Absolute vs. relative assessment

Based on the required characteristics, the main implication is that methods that for which the assessment of one alternative depends on other alternatives being assessed at the same time are not the most adequate for three reasons: first, an absolute rather than a relative assessment is preferred (Requisite 5 in Table 2); second, it should be possible for an organization developing a new chemical or material to assess if it is SSbD, without needing to obtain data on all of its possible competitors; and third, because rank-reversals (e.g., A is better than B if C is not considered, but B is better than A if C is present) (Wang et al., 2009) are better avoided. This rules out using some of the methods identified in the review (Table 2), such as AHP or PROMETHEE II, to the extent they depend on competing alternatives.

As discussed by Dias et al. (2019), even methods that in principle perform an absolute assessment can subtly contain a relative assessment component in the form of a scale normalization, for methods that require all attributes are given on a common scale. If the normalization operation is based on the maximum and/or minimum performances observed in the set of alternatives being compared, as is often the case, then the evaluation becomes relative. This rules out other approaches identified in the review, such as normalization-based Weighted sums, TOPSIS and VIKOR. Therefore, any absolute assessment method must either be exempt from normalization or use a normalization anchored on external references (which is also advocated in the SSbD framework for Step 4).

3.1.2. Input and output scales

Per the third requisite in Table 2, the result should be a rating level (e.g., “A”, “B”, “C” …, or “Green”, “Yellow”, “Red”) or a cardinal value (e.g. 56.78). It is important to note that using numerical levels (e.g., levels 0, 1, 2, 3, or the number of stars in a hotel) does not necessarily correspond to cardinal values if these numerals could be replaced by non-numerical labels. Indeed, a rating level is defined in an ordinal scale that represents only the ranking of the levels (e.g., “A” is better than “B”, or “Green” is better than “Yellow”). To be considered an interval scale, equal differences between levels must correspond to the same meaning concerning what is being measured (e.g., the difference between a “0” and a “1” is as important or as valued as the difference between a “2” and a “3”).

It is possible to convert numeric values (i.e. values on a quantitative interval scale) into ratings and vice versa, by eliciting further information. The conversion of numeric values into ratings requires eliciting the thresholds separating the consecutive levels (Fig. 2). On the reverse direction, conversion of ratings into numeric values requires eliciting a value function (Fig. 3), through a direct or an indirect elicitation protocol (Morton and Fasolo, 2009; Morton, 2018).

The examples provided in the SSbD framework tend to use rating scales, as do most existing frameworks (Caldeira et al., 2022b). The number of levels in other frameworks lies between 3 (e.g., Greenscreen (2018)) and 5 (Caldeira et al., 2022b). In the SSbD framework the number of levels suggested for the different steps also varies between 3 and 5. The option of using rating scales for the SSbD framework presents several advantages:

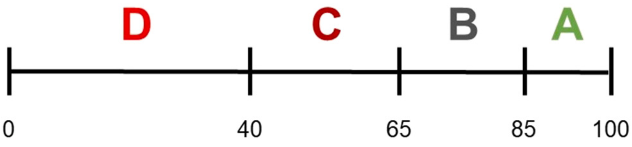


Fig. 2. Converting *numeric* values to rating levels (illustrative example).

- It is well aligned with existing scoring or certification schemes proposed by different entities (Caldeira et al., 2022b, p. 85);
- It lends itself to an easy interpretation, allowing the use of a colour coding, facilitating communication with the public;
- It suits well the association between levels and actions, e.g., level 3 entails full authorization, level 2 entails conditional authorization, etc.;
- It fits assessments based on qualitative properties (e.g., carcinogenic category, being flammable, etc.) rather than numbers, as occurs for the Step 1 hazards-based assessment in the SSbD framework;
- It avoids an illusory perception of precision when reading results;
- It allows a greater stability of the assessment result with regards to some uncertainties.

On the other hand, using numeric values for the SSbD framework also has some advantages:

- It preserves the maximum amount of information in the aggregation step (e.g., it is a B, but very close to an A);
- It allows more discrimination when comparing alternatives;
- It is not difficult to translate results as a rating (Fig. 2).

If the inputs scale for the attributes is quantitative, it is possible to convert it at any moment to a rating scale, but the contrary is considered to be harder. The possible pathways along the aggregation hierarchy can then be, at each stage (Fig. 4):

a) aggregation of rating levels given as an input to provide an output as a rating level.
b) aggregation of numeric values given as an input to provide an output as a rating level;
c) aggregation of numeric values given as an input to provide an output as a numeric value;

3.2. Aggregation at each level of the hierarchy

Aggregation is required at the top of the hierarchy, as well as within each dimension. Different methods can be used for different nodes of the hierarchy (Fig. 1) to aggregate the respective “children” nodes. The SSbD framework already proposes how the aggregation could be performed at some of the nodes in the hierarchy in Fig. 1, whereas for the remaining nodes the aggregation is yet to be defined (see Table 3). Namely, referring back to the possibilities a) to c) in Fig. 4.

3.2.1. Aggregation of rating levels given as an input to provide an output as a rating level

Aggregation based on IF-THEN rules is a possible way to aggregate rating levels. This has the advantages of respecting the qualitative nature of the scale, being simple to read, and implementing the principles that the aggregation shall not allow compensation. In particular, the SSbD framework Requisite 6 and Requisite 7 (Table 2) includes an IF-THEN rule: if the minimum criteria for safety dimension are not met, then the chemical/material cannot be considered as SSbD. The Decision EXpert (DEX) method (Bohanec et al., 2013) implements decision rules following an expert system concept. The Dominance based Rough Set Approach (DRSA) offers a methodology to develop IF-THEN rules from classification examples provided by decision makers (Greco et al., 2016). Such rules will be in the form “Chemical C is rated L3 if it is L3 in attribute 1 and L2 or L3 in attribute 2 and L2 or L3 in attribute 3 and

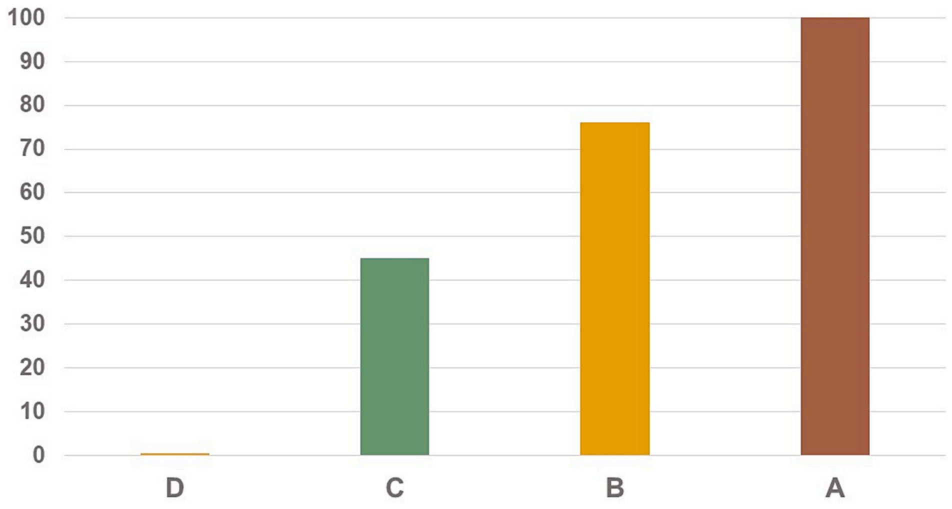


Fig. 3. Converting rating levels to *numeric* values (illustrative example).

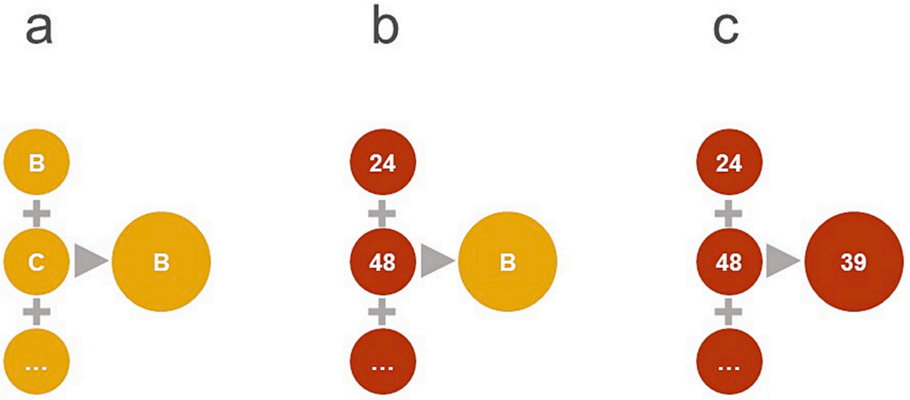


Fig. 4. Aggregation at each level: a) rating to rating; b) cardinal to rating; c) cardinal to cardinal.

Table 3
Proposals in the SSbD framework concerning aggregation.

Attribute	Aggregation proposed in the SSbD framework	Comments
1 Safe	Inputs are rating levels. No proposal has been made for the aggregation, except that a Not-SSbD rating in Step 1 (intrinsic hazard) is a cut-off criterion.	An aggregation based on the rating levels should be defined, as the inputs are given as rating levels.
1.1 Step 1 (intrinsic hazard)	A set of rules based on passing three criteria (conditions) defines whether the chemical is rated Not-SSbD, Level 1, Level 2 or Level 3.	Rating scale with four levels, which matches well the fact that the criteria refer to qualities of the chemicals, rather than quantified properties.
1.2 Step 2 (process)	Aggregation of 5 attributes (aspects), each one assessed using a 5-level rating scale. It is suggested these ratings could be added, and then transformed into a rating on the same scale using thresholds.	The inputs are rating levels, which match well the fact that the criteria refer to qualities of the chemicals, rather than quantified properties. Adding the rating levels is not recommended because this does not match the qualitative nature of the inputs scale, and it allows compensation.
1.3 Step 3 (use phase)	Aggregation of two attributes, each one associated 5 levels with a Pass threshold. Rating level 0 if the chemical fails to pass both criteria; level 1 if it passes the human health criterion, level 2 if it passes both criteria.	The inputs are initially given as cardinals (% above or below safe level), translated into 5 rating levels, and therefore the output could offer a finer discrimination than 3 levels only.
2 Sustainable	A set of rules based on passing four criteria (conditions) defines whether the chemical is rated Level 0, Level 1, …, Level 4. The order of testing defines a lexicographic order of importance for the criteria	The suggestion can be adopted or an option relying less strictly on the pass/fail levels can be devised.
2.1 Toxicity 2.3 Pollution 2.4 Resources	No proposal has been made for the aggregation. The inputs are initially cardinal (% above or below target level), translated into 5 rating levels. Thresholds also define pass/fail criteria.	Two options can be considered: an aggregation based on the ratings or an aggregation based on the initial interval scales.
2.2 CC	(Single indicator: no aggregation needed)	

better than L2 in attribute 4, OR if it is L4 in attribute 1 and L1 in attribute 2 and L1 or L2 in attribute 3 and better than L3 in attribute 4, OR …" (many AND-OR rules may be needed to define each level, but the DRSA approach will bring the number of rules to a minimum).

A second possibility is to use a concordance-discordance voting analogy, as illustrated in Fig. 5 for a situation with 4 levels (which can be adapted to any number of levels). This method can be seen as a qualitative version of the ELECTRE TRI method originally developed by Yu

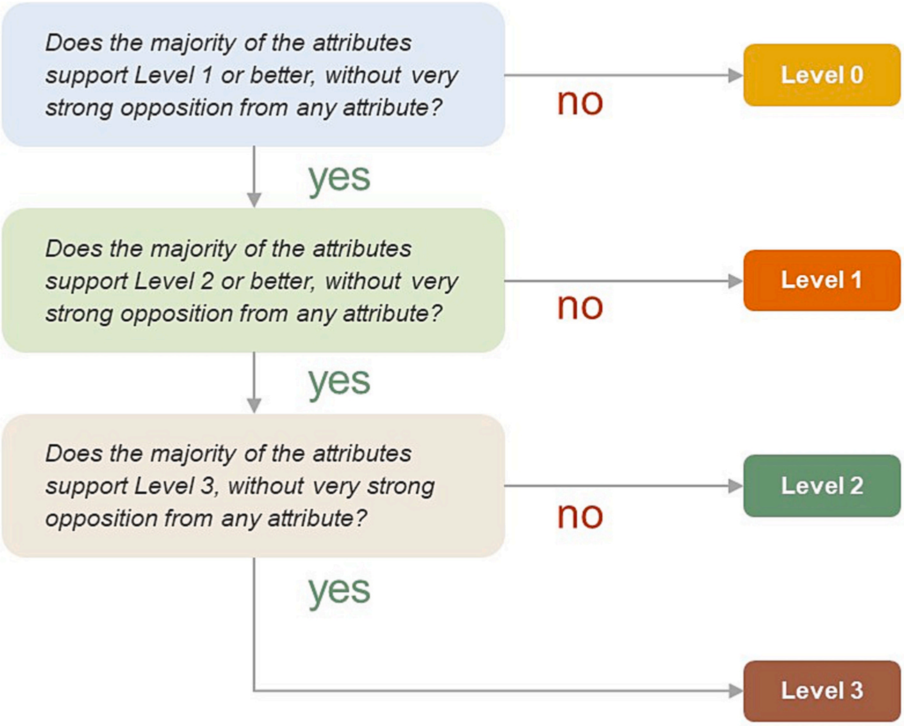


Fig. 5. Flowchart for a concordance-discordance aggregation.

(1992) for cardinal scales (see also Dias et al., 2018; Dias and Mousseau, 2018; Figueira et al., 2013; Govindan and Jepsen, 2016). It requires defining what is considered a sufficient majority, by defining a majority level (e.g., 51 %, 60 %, 2/3) and by defining the number of votes for each attribute (i.e. granting more weight to attributes considered to be more important). This method also allows defining that a very poor performance on some attribute can veto a majority. For instance, it can be defined that having Level 0 on any attribute vetoes a majority supporting Level 3. An illustration of the qualitative version of the ELECTRE TRI method is provided in SI.

Referring back to Table 3, these ways of aggregating rating levels to provide an output rating level clearly applies to node 1.2 (Step 2 - processing), but also can be applied to all other nodes.

3.2.2. *Aggregation of numeric values given as an input to provide an output as a rating level*

The methods that can be used in a) can also be used in this situation, applying first a thresholds-based conversion such as the one illustrated in Fig. 2 to the input values. This conversion is however not necessary. Indeed, the DRSA and DEX methods, mentioned in a) above, can also be applied when inputs are given on as numeric values. Similarly, the concordance-discordance aggregation can also be applied using the ELECTRE TRI method (Figueira et al., 2013; Yu, 1992), which was developed to be a rating method translating multiple numeric scales onto a qualitative rating. Indeed, the second possibility mentioned in a) is a qualitative version of the ELECTRE TRI. FlowSort (Nemery and Lamboray, 2008) is another rating method that can be used. Without excluding other methods, ELECTRE TRI matches perfectly the logic illustrated above (the flowchart in Fig. 5 still applies), and its non-compensatory nature, allowing to model a veto effect absent from other methods, makes it particularly suitable for sustainability assessment (see, e.g., (Dias, Luis.C., 2021), with a step-by-step example).

Referring back to Table 3, these ways of aggregating numbers given as an input to provide an output rating level can be applied to nodes 1.3 (Step 3 - use phase), 2 (Environmental dimension, as a more flexible alternative to the proposed rules), 2.1 (Toxicity), 2.3 (Pollution), and 2.4 (Resources).

3.2.3. *Aggregation of numeric values given as an input to provide an output as a numeric value*

Many possibilities to aggregate these n values $x_1, \ldots, x_n$ into a global value g have been proposed, with methods AHP, MAVT, PROMETHEE or TOPSIS, being often highlighted in literature reviews (e.g., Cinelli et al., 2014; Kumar et al., 2017; Wang et al., 2009). Several methods are undesirable for the SSbD framework, because they allow that a very poor performance on some criterion can be easily compensated in other attributes. This is notoriously the case of the well-known linear additive weighted method (weighted sum method), which sums the product of values x_i by the respective weights w_i ($i = 1, \ldots, n$). This aggregation requires that all attributes are expressed on the same scale, by means of normalization or, even better, by means of converting the original scales into commensurable value or utility scales.

On the opposite side of the compensation vs. non-compensation spectrum, a simple approach is to take the worst value among $x_1, \ldots, x_n$. Such an operation considers all attributes as equally important and requires that all attributes are assessed on the same scale. Another potential disadvantage is that it does not encourage improvements in one attribute above the minimum in the other attributes.

Other methods fall in-between these extremes. In particular, the Ordered Weighted Average (Yager, 1988) is a flexible way of aggregating values given on a common scale, which includes the weighted sum and the minimum as particular cases. The MAUT is another framework that encompasses the weighted sum as a particular case, but includes more complex aggregation formulas (multiplicative, multilinear) where compensation effects are no longer linear (Keeney and Raiffa, 1993). Finally, it is also possible to combine an additive aggregation with IF-THEN rules that prevent a very poor performance to be compensated on another indicator.

Referring back to Table 3, these ways of aggregating numbers can be applied to the same nodes as b), although this would not be recommended for node 1.3 (Step 3 - use phase), since “sister nodes” 1.1 and 1.2 are assessed on a rating scale.

3.3. Consideration of data quality

Requisite 2 in Table 2 stipulates that the evaluation procedure shall take into account the lack of data and data uncertainty. Organizations such as SETAC, the US Dept. of Agriculture, the US EPA, or the EU JRC have proposed several approaches to address this issue (Edelen and Ingwersen, 2018; Lewandowska et al., 2021). Most of these approaches are inspired on the Pedigree Matrix concept from Funtowicz and Ravetz (1990), as proposed by Weidema and Wesnæs (1996). Its adaptation to the LCA area comprises data quality attributes such as reliability, temporal correlation, or geographical correlation, with minor differences between authors and organizations, using a 1–5 ordinal assessment scale. Focusing on inventories, Qin et al. (2020) proposed a Pedigree Matrix for the impact assessment phase. Although not explicitly based on the Pedigree Matrix, EC Environmental Footprint methods (EC, 2021) also use an ordinal 1–5 scale with regards to data quality attributes.

Having rated multiple data qualitative indicators on such 1–5 scales, some frameworks propose to aggregate them (which is in fact a multicriteria aggregation problem on its own, for which most of the considerations in Sections 3.1–3.2 can be relevant). The most common solution is to compute some sort of average (Lewandowska et al., 2021), but this is problematic for two reasons. First, it interprets the "1" to "5" labels as if these were numbers on a cardinal scale. For instance, this assumes the difference between "1" (Excellent) and "2" (Very Good) is the same as the difference between "4" (Fair) and "5" (Poor). A second reason is the compensatory nature of this aggregation. Using a simple average, having data quality (2,2,2,2,2) on five attributes (very good overall) is considered to be as good as (5,1,1,1,2), but this single 5 rating (Poor) might render the overall results quite uncertain.

Other frameworks have suggested to translate the Pedigree Matrix into probability distributions, either based on existing databases (Ciroth et al., 2016), or based on elicited expert judgment (Qin et al., 2020). Using appropriate uncertainty analysis simulation software, it is then possible to obtain probability distributions for results of interest. However, Qin et al. (2020)'s survey to 47 LCA practitioners shows lack of consensus concerning the use of the Pedigree Matrix, and Edelen and Ingwersen (2018)'s review found authors claiming there is no sound justification for creating probability distributions from data quality assessments.

Data quality assessments can also be used to limit the maximum rating that can be attributed to a chemical. An example is the GreenScreen® method for safety assessment (GreenScreen, 2018). This method assesses chemicals to provide a qualitative rating on a scale from Benchmark-1 to Benchmark-4. In parallel, it defines data quality conditions to reach these levels, meaning that a Benchmark-4 chemical might see its rating lowered to Benchmark-3, or even less, due to not meeting the data quality criteria for a higher benchmark.

More generally, a US National Academies NRC (National Research Council, 2014) defines different ways to cope with uncertainty:

- Using only known best estimates, excluding (not assessing) chemicals with critical data missing.
- Performing uncertainty downgrades (as occurs in GreenScreen®), thereby punishing alternatives with poor data quality, which is deemed by the NRC to be counter-productive.
- Performing quantitative uncertainty analyses, based on ranges or probability distributions, which NRC deems might be sufficient for some comparative assessments.
- Remaining neutral about uncertainty and missing data, noting the presence of uncertainty and missing data but not excluding the alternative, which is the option the NRC considers better aligned with the nature of their framework.

Considering the SSbD framework, aiming at guiding innovation at early design stages, one should consider the purpose of the assessment. While innovating, often through a trial-and-error process, many data might be missing or be highly uncertain. As the innovation process progresses, some options are discarded, more investment in data gathering occurs, and uncertainties tend to decrease. In these settings, excluding chemicals for which good data does not exist yet, or punishing such chemicals with a lower rating does not seem sensible for the innovators. Therefore, they can remain neutral about uncertainty and missing data, noting the presence of uncertainty, as advocated by the NRC. The innovation team might refer to the obtained rating as the "Estimated SSbD rating", which can be compared with the sought rating as a driver for further innovation and data collection efforts. A second possibility is to perform quantitative uncertainty analyses that will indicate the probability distribution for the chemical's rating. This requires some process for estimating the input attributes distributions. The innovation team might refer to the "Most likely SSbD rating", "Median SSbD rating", "Minimum assured SSbD rating", or even to an "SSbD rating 95% confidence interval" to guide their decisions.

A different assessment perspective is that of certification, possibly needed for communicating an SSbD rating to regulators, customers, etc. The issue of data quality then becomes much more important. Remaining neutral about uncertainty and missing data might risk rewarding lack of data. The possibility of performing quantitative uncertainty analyses based on justifiable probability distributions remains open, but in a certification perspective it should provide a conservative rating, such as the minimum or the rating corresponding to, say, the 5th or 10th percentile of the output distribution. The certification might then read, for instance "Certified conservative SSbD rating estimate". Such conservative estimate might be computable by combining suitably conservative estimates for the inputs, even when distributions are not available. Finally, if the overall data quality (as assessed by some adequate multiattribute aggregation method) is considered to be sufficient, a "Certified SSbD rating" could be recognized. Apart from these labels, all other situations would be considered a "Non-certified SSbD rating estimate".

4. Conclusions

This article aimed at identifying the most relevant characteristics of the SSbD evaluation framework concerning the possibility of aggregating safety and sustainability assessments for chemicals and materials, as well as identifying MCDA methods that can potentially be applied.

The strategy of aggregating rating levels given as an input to provide an output as a rating level offers many interesting advantages. In addition to the advantages of expressing the result as a qualitative rating (mentioned in Section 3.1), it is easy to communicate, either by means of IF-THEN rules, or by means of concordance-discordance logic follows a flowchart (Fig. 5) resembling the Step 1 and Step 3 safety assessment suggested in the SSbD framework. Moreover, this facilitates setting conditions that prevent poor performance on one indicator to be compensated on other indicators.

The strategy consisting in an aggregation of numeric values given as an input to provide an output as a rating level also has the advantages of expressing the result as a qualitative rating. Among the possibilities considered, methods implementing IF-THEN rules or ELECTRE TRI stand out as being perfectly aligned with the previous one, which leads to a more harmonious evaluation framework of if both strategies are used in different nodes of the hierarchy.

The strategy of aggregating numeric values given as an input to provide a numeric value as an output, in turn, has the advantage of preserving information assessed on cardinal scales, but cannot be used in nodes of the hierarchy where the inputs are rating levels, hindering the overall harmony of the framework. Moreover, some methods for this strategy might unduly allow compensation of poor performance and may also require the inputs to be given on a common scale.

Considering data quality, one should consider what the purpose of the assessment is. For internal innovation processes, an innovation team might use their best estimates and refer to the obtained rating as the

"Estimated SSbD rating", which can be compared with the sought rating as a driver for further innovation and data collection efforts. A second possibility is to perform quantitative uncertainty analyses that will indicate the "Most likely SSbD rating", the "Median SSbD rating", the "Minimum assured SSbD rating", or even an "SSbD rating 95% confidence interval" to guide innovation decisions. In a perspective of certification, possibly needed for communicating an SSbD rating to regulators, customers, etc., quantitative uncertainty analyses based on justifiable probability distributions can provide a "Certified conservative SSbD rating estimate", unless the overall data quality (as assessed by some adequate multiattribute aggregation method) is considered to be sufficient to warrant a "Certified SSbD rating".

In any case, the use of multiattribute aggregation does not rule out a richer dashboard presenting not only the overall aggregate result, but also the results obtained in other levels of the hierarchy. Such complementary information is important to understand the strengths and weaknesses that an aggregate result inevitably might hide.

A limitation of the present work is that the current EC SSbB framework is expected to be subject to further refinements and to be perfected upon performing several case studies, and no specific case study was addressed here. Future case studies may allow trying out all these possibilities, aiming at finding the most satisfactory one(s). Another limitation of this work is not addressing specifically the issue of choosing a weighting vector, or setting the aggregation methods rules or parameters. Future case studies will also offer a testbed for different approaches. Finally, a limitation of the literature review is its focus on keywords associated with multiattribute aggregation, which might have missed some studies performing some kind of MCDA without explicitly mentioning it in the title, abstract or keywords.

CRediT authorship contribution statement

Luis C. Dias: Conceptualization, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. **Carla Caldeira:** Conceptualization, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. **Serenella Sala:** Conceptualization, Supervision, Writing – review & editing.

Declaration of competing interest

None.

Data availability

Data used in the study are available in the manuscript

Acknowledgments

The work presented in this paper was developed in the context of the Administrative Arrangement "Support Criteria for Safe and Sustainable-by-Design advanced materials and chemicals (SSBDCHEM)", JRC 36058 / DG RTD LC-01671974, between the European Commission's Directorate General for Research and Innovation (DG RTD) and the Joint Research Centre (JRC).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2023.169599.

References

Abdel-Basset, M., Gamal, A., Chakrabortty, R.K., Ryan, M.J., 2021. Evaluation of sustainable hydrogen production options using an advanced hybrid MCDM approach: a case study. Int. J. Hydrog. Energy 46 (5), 4567–4591. https://doi.org/10.1016/j.ijhydene.2020.10.232.

Acar, C., Beskese, A., Temur, G.T., 2018. Sustainability analysis of different hydrogen production options using hesitant fuzzy AHP. Int. J. Hydrog. Energy 43 (39), 18059–18076. https://doi.org/10.1016/j.ijhydene.2018.08.024.

Acar, C., Beskese, A., Temur, G.T., 2022. Comparative fuel cell sustainability assessment with a novel approach. Int. J. Hydrog. Energy 47 (1), 575–594. https://doi.org/10.1016/j.ijhydene.2021.10.034.

Alkhatib, I.I.I., Khalifa, O., Bahamon, D., Abu-Zahra, M.R.M., Vega, L.F., 2021. Sustainability criteria as a game changer in the search for hybrid solvents for CO2 and H2S removal. Sep. Purif. Technol. 277, 119516 https://doi.org/10.1016/j.seppur.2021.119516.

Banimostafa, A., Papadokonstantakis, S., Hungerbühler, K., 2012. Evaluation of EHS hazard and sustainability metrics during early process design stages using principal component analysis. Process Saf. Environ. Prot. 90 (1), 8–26. https://doi.org/10.1016/j.psep.2011.06.014.

Bohanec, M., Žnidaršič, M., Rajkovič, V., Bratko, I., Zupan, B., 2013. DEX methodology: three decades of qualitative multi-attribute modeling. Informatica 37 (1), 49–54.

Bouyssou, D., 1993. Décision Multicritère ou aide multicritère. Newsletter of the European working group "multicriteria aid for decisions". Series 2, Spring, 1–2.

Brans, J.P., Vincke, P., Mareschal, B., 1986. How to select and how to rank projects: the Promethee method. Eur. J. Oper. Res. 24 (2), 228–238. https://doi.org/10.1016/0377-2217(86)90044-5.

Caldeira, C., Farcal, L., Tosches, D., Amelio, A., Rasmussen, K., Rauscher, H., Riego Sintes, J., Sala, S., 2022a. Safe and Sustainable by Design chemicals and materials - Framework for the definition of criteria and evaluation procedure for chemicals and materials. EUR 31100 EN. Publications Office of the European Union.

Caldeira, C., Farcal, L., Moretti, C., Mancini, L., Rauscher, H., Rasmussen, K., Riego Sintes, J., Sala, S., 2022b. Safe and Sustainable by Design Chemicals and Materials - Review of Safety and Sustainability Dimensions, Aspects, Methods, Indicators, and Tools. EUR 30991. Publications Office of the European Union.

Cinelli, M., Coles, S.R., Kirwan, K., 2014. Analysis of the potentials of multi criteria decision analysis methods to conduct sustainability assessment. Ecol. Indic. 46, 138–148. https://doi.org/10.1016/j.ecolind.2014.06.011.

Cinelli, M., Kadziński, M., Gonzalez, M., Słowiński, R., 2020. How to support the application of multiple criteria decision analysis? Let us start with a comprehensive taxonomy. Omega 96, 102261. https://doi.org/10.1016/j.omega.2020.102261.

Cinelli, M., Koffler, C., Askham, C., Amadei, A., Arendt, R., Bachmann, T. M., Barros, B., Bjørn, A., Dias, L. C., Laurent, A., Motoshita, M., Rupcic, L., Sala, S., Santos, J., Scherer, L., Steen, B., & Yokoi, R. (2022). Criteria used to review weighting methods as part of the UN environment life cycle Initiative's global guidance on environmental life cycle impact assessment indicators (GLAM) project. Abstract for the SETAC Europe annual meeting, 15-19 may.

Ciroth, A., Muller, S., Weidema, B., Lesage, P., 2016. Empirically based uncertainty factors for the pedigree matrix in ecoinvent. Int. J. Life Cycle Assess. 21 (9), 1338–1348. https://doi.org/10.1007/s11367-013-0670-5.

Colorni, A., Tsoukiàs, A., 2021. Rating or sorting: terminology matters. J. Multi-Criteria Decis. Anal. 28 (3–4), 131–133. https://doi.org/10.1002/mcda.1733.

Crivellari, A., Casson Moreno, V., Cozzani, V., Dincer, I., 2021. Multi-criteria sustainability assessment of potential methanol production processes. J. Clean. Prod. 293, 126226 https://doi.org/10.1016/j.jclepro.2021.126226.

Dias, Luis.C., 2021. Sustainability assessment using the ELECTRE TRI multicriteria sorting method. In: Ren, J. (Ed.), Methods in Sustainability Science. Elsevier, pp. 197–214. https://doi.org/10.1016/C2020-0-00430-5.

Dias, Luis C., Mousseau, V., 2018. Eliciting multi-criteria preferences: ELECTRE models. In: Dias, Luis C., Morton, A., Quigley, J. (Eds.), Elicitation - The Science and Art of Structuring Judgement. Springer, pp. 349–375. https://doi.org/10.1007/978-3-319-65052-4_14.

Dias, Luis C., Antunes, C.H., Dantas, G., de Castro, N., Zamboni, L., 2018. A multi-criteria approach to sort and rank policies based on Delphi qualitative assessments and ELECTRE TRI: the case of smart grids in Brazil. Omega 76, 100–111. https://doi.org/10.1016/j.omega.2017.04.004.

Dias, Luis C., Freire, F., Geldermann, J., 2019. Perspectives on multi-criteria decision analysis and life-cycle assessment. In: Doumpos, M., Figueira, J.R., Greco, S., Zopounidis, C. (Eds.), New Perspectives in Multiple Criteria Decision Making. Springer, pp. 315–329. https://doi.org/10.1007/978-3-030-11482-4_12.

Diaz-Balteiro, L., González-Pachón, J., Romero, C., 2017. Measuring systems sustainability with multi-criteria methods: a critical review. Eur. J. Oper. Res. 258 (2), 607–616. https://doi.org/10.1016/j.ejor.2016.08.075.

Dinh, L.T.T., Guo, Y., Mannan, M.S., 2009. Sustainability evaluation of biodiesel production using multicriteria decision-making. Environ. Prog. Sustain. Energy 28 (1), 38–46. https://doi.org/10.1002/ep.10335.

EC, 2019. Communication from the commission to the European Parliament, the council, the European economic and social committee and the Committee of the Regions. The European Green Deal. COM 2019, 640.

EC, 2020. Communication from the commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. Chem. Strategy Sustain. Towards a Toxic-Free Environ. COM 2020, 667.

EC (2021). Commission recommendation of 16.12.2021 on the use of the Environmental Footprint methods to measure and communicate the life cycle environmental performance of products and organisations. 9332.

EC. (2022). Commission recommendation of 8.12.2022 establishing a European assessment framework for 'safe and sustainable by design' chemicals and materials, C(2022) 8854 final.

Edelen, A., Ingwersen, W.W., 2018. The creation, management, and use of data quality information for life cycle assessment. Int. J. Life Cycle Assess. 23 (4), 759–772. https://doi.org/10.1007/s11367-017-1348-1.

Ehrgott, M., Stewart, T.J. (Eds.), 2010. Multiple Criteria Decision Making for Sustainable Energy and Transportation Systems. Springer.

Figueira, J., Greco, S., Roy, B., Slowinski, R., 2013. An overview of ELECTRE methods and their recent extensions. J. Multi-Criteria Decis. Anal. 20, 61–85. http://www3.interscience.wiley.com/journal/5725/home?SRETRY=0.

Forman, E.H., Gass, S.I., 2001. The analytic hierarchy process—an exposition. Oper. Res. 49 (4), 469–486. https://doi.org/10.1287/opre.49.4.469.11231.

Funtowicz, S.O., Ravetz, J.R., 1990. Uncertainty and Quality in Science for Policy. Springer Science & Business Media.

Goedkoop, M., Heijungs, R., Huijbregts, M., De Schryver, A., Struijs, J., Van Zelm, R., 2009. ReCiPe 2008. Ministerie van Volkshuisvesting, Ruimtelijke ordening en Milieubeheer doi:10.029/2003JD004283, 126 pages.

Govindan, K., Jepsen, M.B., 2016. ELECTRE: a comprehensive literature review on methodologies and applications. Eur. J. Oper. Res. 250 (1), 1–29. https://doi.org/10.1016/j.ejor.2015.07.019.

Greco, S., Matarazzo, B., Słowiński, R., 2016. Decision rule approach. In: Multiple Criteria Decision Analysis: State of the Art Surveys, pp. 497–552. https://doi.org/10.1007/978-1-4939-3094-4_13.

GreenScreen, 2018. GreenScreen method for Safer Chemicals. https://www.greenscreenchemicals.org/learn/full-greenscreen-method (accessed on 12/12/2022.

Huang, I.B., Keisler, J., Linkov, I., 2011. Multi-criteria decision analysis in environmental sciences: ten years of applications and trends. Sci. Total Environ. 409 (19), 3578–3594. https://doi.org/10.1016/j.scitotenv.2011.06.022.

Huijbregts, M., Steinmann, Z.J.N., Elshout, P.M.F.M., Stam, G., Verones, F., Vieira, M., Zijp, M., Hollander, A., van Zelm, R., 2016. ReCiPe 2016 - a harmonized life cycle impact assessment method at midpoint and endpoint level. In: Report I: Characterization. National Institute for Public Health and the Environment, p. 194.

Huysveld, S., Taelman, S.E., Hackenhaar, I.C., Zanchi, L., Kujanpää, L., Harmens, R., Alvarenga, R., 2021. Critical evaluation of sustainability integration approaches. D1.5 Orienting project. https://orienting.eu/publications/page/2/.

Ibáñez-Forés, V., Bovea, M.D., Pérez-Belis, V., 2014. A holistic review of applied methodologies for assessing and selecting the optimal technological alternative from a sustainability perspective. J. Clean. Prod. 70, 259–281. https://doi.org/10.1016/j.jclepro.2014.01.082.

Iranfar, S., Karbala, M.M., Shahsavari, M.H., Vandeginste, V., 2023. Prioritization of habitat construction materials on Mars based on multi-criteria decision-making. J. Build. Eng. 66, 105864 https://doi.org/10.1016/j.jobe.2023.105864.

Janošovský, J., Boháčiková, V., Kraviarová, D., Variny, M., 2022a. Multi-criteria decision analysis of steam reforming for hydrogen production. Energy Convers. Manag. 263, 115722 https://doi.org/10.1016/j.enconman.2022.115722.

Janošovský, J., Boháčiková, V., Rosa, I., Vincent, G., Labovská, Z., Jelemenský, Ľ., 2022b. Recognition of process safety position in multiple-criteria decision analysis in process design. Chem. Eng. Trans. 90, 775–780.

Jia, X., Li, Z., Wang, F., Qian, Y., 2016. Integrated sustainability assessment for chemical processes. Clean Techn. Environ. Policy 18 (5), 1295–1306. https://doi.org/10.1007/s10098-015-1075-x.

Keeney, R.L., 2006. Using preferences for multi-attributed alternatives. J. Multi-Criteria Decis. Anal. 14, 169–174. https://doi.org/10.1002/mcda.

Keeney, R.L., Raiffa, H., 1993. Decisions with Multiple Objectives–Preferences and Value Tradeoffs. Cambridge University Press.

Kumar, A., Sah, B., Singh, A.R., Deng, Y., He, X., Kumar, P., Bansal, R.C., 2017. A review of multi criteria decision making (MCDM) towards sustainable renewable energy development. Renew. Sust. Energ. Rev. 69, 596–609. https://doi.org/10.1016/j.rser.2016.11.191.

Lewandowska, A., Joachimiak-Lechman, K., Kurczewski, P., 2021. A dataset quality assessment—an insight and discussion on selected elements of environmental footprints methodology. Energies 14 (16), 5004. https://doi.org/10.3390/en14165004.

Li, X., Zanwar, A., Jayswal, A., Lou, H.H., Huang, Y., 2011. Incorporating exergy analysis and inherent safety analysis for sustainability assessment of biofuels. Ind. Eng. Chem. Res. 50 (5), 2981–2993. https://doi.org/10.1021/ie101660q.

Liew, W., Hassim, M., Ng, D., 2013. Screening of sustainable biodiesel production pathways during process Research and Development (R&D) stage using fuzzy optimization. Chem. Eng. Trans. 35, 1075–1080. https://doi.org/10.3303/CET1335179.

Limleamthong, P., Gonzalez-Miquel, M., Papadokonstantakis, S., Papadopoulos, A.I., Seferlis, P., Guillén-Gosálbez, G., 2016. Multi-criteria screening of chemicals considering thermodynamic and life cycle assessment metrics via data envelopment analysis: application to CO 2 capture. Green Chem. 18 (24), 6468–6481. https://doi.org/10.1039/C6GC01696K.

Lindfors, A., 2021. Assessing sustainability with multi-criteria methods: a methodologically focused literature review. Environ. Sustain. Indicators 12, 100149. https://doi.org/10.1016/j.indic.2021.100149.

Linkov, I., Moberg, E., 2012. Multi-Criteria Decision Analysis: Environmental Applications and Case Studies. CRC Press.

Linkov, I., Moberg, E., Trump, B.D., Yatsalo, B., Keisler, J.M., 2020. Multi-Criteria Decision Analysis: Case Studies in Engineering and the Environment. CRC Press.

Marttunen, M., Suomalainen, M., 2005. Participatory and multiobjective development of WaterCourse regulation creation of regulation alternatives from stakeholders ` preferences. J. Multi-Criteria Decis. Anal. 49 (2005), 29–49.

Morales, M., Ataman, M., Badr, S., Linster, S., Kourlimpinis, I., Papadokonstantakis, S., Hatzimanikatis, V., Hungerbühler, K., 2016. Sustainability assessment of succinic acid production technologies from biomass using metabolic engineering. Energy Environ. Sci. 9 (9), 2794–2805. https://doi.org/10.1039/C6EE00634E.

Moretti, L., Di Mascio, P., Bellagamba, S., 2017. Environmental, human health and socio-economic effects of cement powders: the multicriteria analysis as decisional methodology. Int. J. Environ. Res. Public Health 14 (6), 645. https://doi.org/10.3390/ijerph14060645.

Morton, Alec, 2018. Multiattribute value elicitation. In: Dias, Luis C., Morton, A., Quigley, J. (Eds.), Elicitation: The Science and Art of Structuring Judgement. Springer International Publishing, pp. 287–311. https://doi.org/10.1007/978-3-319-65052-4_12.

Morton, A., Fasolo, B., 2009. Behavioural decision theory for multi-criteria decision analysis: a guided tour. J. Oper. Res. Soc. 60 (2), 268–275. https://doi.org/10.1057/palgrave.jors.2602550.

Munda, G., 2004. Social multi-criteria evaluation: methodological foundations and operational consequences. Eur. J. Oper. Res. 158 (3), 662–677. https://doi.org/10.1016/S0377-2217(03)00369-2.

Munda, G., 2005. Multiple criteria decision analysis and sustainable development. In: Figueira, J., Greco, S., Ehrgott, M. (Eds.), Multiple Criteria Decision Analysis: State of the Art Surveys. Springer.

Mustajoki, J., Hämäläinen, R.P., Marttunen, M., 2004. Participatory multicriteria decision analysis with web-HIPRE: a case of lake regulation policy. Environ. Model Softw. 19 (6), 537–547. https://doi.org/10.1016/j.envsoft.2003.07.002.

Narayanan, D., Zhang, Y., Mannan, M.S., 2007. Engineering for sustainable development (ESD) in bio-diesel production. Process. Saf. Environ. Prot. 85, 349–359. https://doi.org/10.1205/psep07016.

National Research Council, 2014. A Framework to Guide Selection of Chemical Alternatives. The National Academies Press. https://doi.org/10.17226/18872.

Nemery, P., Lamboray, C., 2008. FlowSort: a flow-based sorting method with limiting or central profiles. TOP 16 (1), 90–113. https://doi.org/10.1007/s11750-007-0036-x.

Pizzol, M., Laurent, A., Sala, S., Weidema, B., Verones, F., Koffler, C., 2017. Normalisation and weighting in life cycle assessment: quo vadis? Int. J. Life Cycle Assess. 22 (6), 853–866. https://doi.org/10.1007/s11367-016-1199-1.

Pizzol, L., Livieri, A., Salieri, B., Farcal, L., Soeteman-Hernández, L.G., Rauscher, H., Zabeo, A., Blosi, M., Costa, A.L., Peijnenburg, W., Stoycheva, S., Hunt, N., López-Tendero, M.J., Salgado, C., Reinosa, J.J., Fernández, J.F., Hristozov, D., Hristozov, D., 2023. Screening level approach to support companies in making safe and sustainable by design decisions at the early stages of innovation. Cleaner Environ. Syst. 10, 100132.

Posada, J.A., Patel, A.D., Roes, A., Blok, K., Faaij, A.P.C., Patel, M.K., 2013. Potential of bioethanol as a chemical building block for biorefineries: preliminary sustainability assessment of 12 bioethanol-based products. Bioresour. Technol. 135, 490–499. https://doi.org/10.1016/j.biortech.2012.09.058.

Purker, S., Lalyer, C.R., Giese, B., 2023. Decision support for selection of new materials considering socio-economic and broader environmental aspects. Sustain. Prod. Consump. 39, 438–450. https://doi.org/10.1016/j.spc.2023.05.032.

Qin, Y., Cucurachi, S., Suh, S., 2020. Perceived uncertainties of characterization in LCA: a survey. Int. J. Life Cycle Assess. 25 (9), 1846–1858. https://doi.org/10.1007/s11367-020-01787-9.

Roesch, A., Sala, S., Jungbluth, N., 2020. Normalization and weighting: the open challenge in LCA. Int. J. Life Cycle Assess. 25 (9), 1859–1865. https://doi.org/10.1007/s11367-020-01790-0.

Roy, B., 1996. Multicriteria Methodology for Decision Aiding. Kluwer Academic.

Saaty, T.L., 2005. The analytic hierarchy and analytic network processes for the measurement of intangible criteria and for decision-making. In: Greco, S., Ehrgott, M., Figueira, J.R. (Eds.), Multiple Criteria Decision Analysis: State of the Art Surveys, vol. 78. Springer-Verlag, pp. 345–405. https://doi.org/10.1007/0-387-23081-5-9.

Saavalainen, P., Kabra, S., Turpeinen, E., Oravisjärvi, K., Yadav, G.D., Keiski, R.L., Pongrácz, E., 2015. Sustainability assessment of chemical processes: evaluation of three synthesis routes of DMC. J. Chemother. 2015, 1–12. https://doi.org/10.1155/2015/402315.

Sala, S., Cerutti, A.K., Pant, R., 2018. Development of a Weighting Approach for the Environmental Footprint. JRC Technical Report.

Samani, P., Mendes, A., Leal, V., Miranda Guedes, J., Correia, N., 2015. A sustainability assessment of advanced materials for novel housing solutions. Build. Environ. 92, 182–191. https://doi.org/10.1016/j.buildenv.2015.04.012.

Seker, S., Aydin, N., 2022. Assessment of hydrogen production methods via integrated MCDM approach under uncertainty. Int. J. Hydrog. Energy 47 (5), 3171–3184. https://doi.org/10.1016/j.ijhydene.2021.07.232.

Shi, X., Jungwirth, S., Akin, M., Wright, R., Fay, L., Veneziano, D.A., Zhang, Y., Gong, J., Ye, Z., 2014. Evaluating snow and ice control Chemicals for Environmentally Sustainable Highway Maintenance Operations. J. Transp. Eng. 140 (11) https://doi.org/10.1061/(ASCE)TE.1943-5436.0000709.

Simanovska, J., Grigale-Soročina, Z., 2016. Design for sustainability approach in product development – a case study using innovative nail polishes developed by a small enterprise. Proc. Est. Acad. Sci. 65 (3), 297. https://doi.org/10.3176/proc.2016.3.11.

Stoycheva, S., Zabeo, A., Pizzol, L., Hristozov, D., 2022. Socio-economic life cycle-based framework for safe and sustainable Design of Engineered Nanomaterials and Nano-Enabled Products. Sustainability 14 (9), 5734. https://doi.org/10.3390/su14095734.

Sun, S., Fu, L., Yang, A., Shen, W., 2023. An intensified energy-saving architecture for side-stream extractive distillation of four-azeotrope mixtures considering economic, environmental and safety criteria simultaneously. Sep. Purif. Technol. 310, 123132 https://doi.org/10.1016/j.seppur.2023.123132.

Tahmid, M., Syeda, S.R., 2023. A framework for integrating safety and environmental impact in the conceptual design of chemical processes. Pure Appl. Chem. 95 (3), 301–314. https://doi.org/10.1515/pac-2022-1120.

van Dijk, J., Flerlage, H., Beijer, S., Slootweg, J.C., van Wezel, A.P., 2022. Safe and sustainable by design: a computer-based approach to redesign chemicals for reduced environmental hazards. Chemosphere 296, 134050. https://doi.org/10.1016/j.chemosphere.2022.134050.

Wang, J., Jing, Y., Zhang, C., Zhao, J., 2009. Review on multi-criteria decision analysis aid in sustainable energy. Renew. Sust. Energ. Rev. 13, 2263–2278. https://doi.org/10.1016/j.rser.2009.06.021.

Weidema, B.P., Wesnæs, M.S., 1996. Data quality management for life cycle inventories—an example of using data quality indicators. J. Clean. Prod. 4 (3–4), 167–174. https://doi.org/10.1016/S0959-6526(96)00043-1.

Yager, R.R., 1988. On ordered weighted averaging aggregation operators in multicriteria decisionmaking. IEEE Trans. Syst. Man Cybern. 18 (1), 183–190. https://doi.org/10.1109/21.87068.

Yu, W., 1992. ELECTRE TRI: Aspects méthodologiques et manuel d'utilisation. Document du LAMSADE No. 74. Université de Paris–Dauphine.

Zanghelini, G.M., Cherubini, E., Soares, S.R., 2018. How multi-criteria decision analysis (MCDA) is aiding life cycle assessment (LCA) in results interpretation. J. Clean. Prod. 172, 609–622. https://doi.org/10.1016/J.JCLEPRO.2017.10.230.