Selecting indicators for measuring progress towards sustainable development goals at the global, national and corporate levels

Gebara, Caroline H.; Thammaraksa, Chonlawan; Hauschild, Michael; Laurent, Alexis

Published in:
Sustainable Production and Consumption

Link to article, DOI:
10.1016/j.spc.2023.12.004

Publication date:
2024

Document Version
Publisher's PDF, also known as Version of record

Sustainable Production and Consumption 44 (2024) 151–165

Contents lists available at ScienceDirect

Sustainable Production and Consumption

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

Selecting indicators for measuring progress towards sustainable development goals at the global, national and corporate levels

Caroline H. Gebara *, Chonlawan Thammaraksa, Michael Hauschild, Alexis Laurent

Quantitative Sustainability Assessment Section, Department of Environmental and Resources Engineering, Technical University of Denmark, Building 115, 2800 Kgs. Lyngby, Denmark

ARTICLE INFO

Editor: Prof. Shabbir Gheewala

Keywords:
Sustainable development goals
Sustainability indicators
Sustainability reporting
Selection criteria
Corporate sustainability

ABSTRACT

The Sustainable Development Goals (SDGs) guide towards sustainability across a broad palette of aspects. Since their adoption by the United Nations in 2015, many organizations have started using them to relate their own sustainability performance. However, there exists no consensus nor science-based methods for performing such assessments, hence bringing the risk of cherry-picking and arbitrary target-setting among the 17 goals and potentially compromising the robustness, usefulness, and credibility of SDG assessment. To address this, we propose a framework that enables to systematically evaluate or select indicators for assessing and monitoring SDG performance at different scales. A critical review of sustainability indicator selection criteria was performed, forming the basis for recommendations of existing criteria. Two separate applications of the framework are operationalized, namely the evaluation of existing indicator sets and the selection of new indicators, respectively. The guide is tested and demonstrated on two different proof-of-concept cases representing the global level and the company level. The results demonstrate that the proposed framework allows decision-makers to systematically evaluate and select SDG indicators for different scopes, offering stronger scientifically founded indicator sets and reducing unintentional bias. Furthermore, the framework still provides the flexibility of tailoring indicators in the selection process, allowing the user to freely select the indicators to be tested. The framework and criteria are central inputs for the indicator selection phase and represent an essential step towards a full-fledged, consistent framework for SDG performance assessments.

1. Introduction

The Sustainable Development Goals (SDGs) framework has gained a lot of attention since its adoption in 2015, when 193 member states of the United Nations (UN) agreed to work towards achieving global sustainable development (United Nations, 2015). The framework comprises 17 goals and 169 targets covering societal and environmental aspects for the world to achieve. Since then, the world has continued facing increasing challenges with regard to sustaining life on Earth as we know it. While the economy continues to grow globally, many people still live in severe poverty without access to basic needs (Millward-Hopkins et al., 2020; Sullivan and Hickel, 2023). The world is already facing tremendous impacts of climate change, causing severe drought, floods, and extreme weather (IPCC, 2022a; Kim et al., 2014), while the net GHG emissions in the previous decade have been the largest ever, intensifying even further the risk for those events (IPCC, 2022a) and destabilizing the lives of many even further (Field et al., 2012). These increasing pressures have motivated the international society to agree on a common limit for climate change with the view of keeping the global temperature increase below 2 degrees Celsius (i.e., Paris Agreement) (UNFCCC, 2015) and recently, an agreement to protect 30 % land and sea area to halt biodiversity loss (i.e., COP15 on global biodiversity) (United Nations, 2022). To achieve these targets, it is crucial that all sectors deliver towards the global challenges, including sub-national and trans-national actors such as businesses and public and private organizations (IPCC, 2022a, 2022b).

To measure the progress towards each of the 17 goals and their targets, the UN has proposed indicators to monitor SDG performance at global and national level (United Nations, 2019). Furthermore, a growing number of assessments at global, member state, and sub-national levels (e.g., companies) monitor progress towards the SDGs (Allen et al., 2018; Sachs et al., 2023). However, the current set of UN indicators is the result of an international policy process, and thus many indicators are not suitable across all countries, nor are they all relevant

* Corresponding author.
E-mail address: cahege@dtu.dk (C.H. Gebara).

https://doi.org/10.1016/j.spc.2023.12.004
Received 4 August 2023; Received in revised form 1 December 2023; Accepted 3 December 2023
Available online 7 December 2023

at sub-national scale. This has resulted in practitioners at national and company scales making arbitrary choices on indicators to fit their contexts without much scientific foundation (Heras-Saizarbitoria et al., 2022; Pizzi et al., 2021). Clearly, in the absence of a formalized and official approach, practitioners can freely decide how they measure progress towards the goals, as there is currently no consensus about how to do this (Hák et al., 2016). This has led to a series of issues. Firstly, practitioners have started 'cherry-picking' SDGs or indicators to measure their progress against, often selecting the SDGs that favor their business while leaving others out (Heras-Saizarbitoria et al., 2022; Lu et al., 2021). This has been criticized by many, arguing that it leads to burden-shifting and, in many cases, 'SDG-washing' when particularly negative impacts on SDGs are overlooked (Heras-Saizarbitoria et al., 2022; Lu et al., 2021; van Zanten and van Tulder, 2021). Secondly, indicators proposed at the global and national levels are generally poorly adapted to the organizational level (e.g., companies). Therefore, there is a need for guidance that can provide decision-makers with science-based and informed support when selecting indicators and assessing SDG performance (Allen et al., 2021; Hák et al., 2016; Janoušková et al., 2018).

Prior research efforts have examined the use of selection criteria for sustainability indicators through extensive literature reviews (Bonisoli et al., 2018; Niemeijer and de Groot, 2008; Pires et al., 2020). However, in addition to not being specific to SDGs, these investigations often exhibit limitations, as they tend to be confined to specific sectors or facets of sustainability, have limited coverage of the literature, or focus on individual indicators omitting criteria for evaluating indicator sets as a whole (needed for guaranteeing cumulative exhaustivity of the indicators). Other studies have developed indicator sets using literature reviews and expert opinions to feed into the selection process (Allen et al., 2020; Cagno et al., 2019; Mengistu and Panizzolo, 2023). Yet, like for sustainability indicators at large, these efforts predominantly cater to specific contexts, such as sectoral or national assessments, where the indicator selections are case-specific and are difficult to generalize. As the SDG framework is implemented across various sectors and organizations, all working towards the common goal of monitoring progress across the 17 SDGs, there is a pressing need for standardized, cross-sectoral approaches to indicator selection, thus enabling more consistent, effective, and comparable monitoring of the goals. This need is aligned with previous calls, like Hák et al., (2016), who highlighted the need for conceptually and methodologically refined approaches to achieve uniformity in SDG performance monitoring instead of focusing on the mere production of isolated, new statistics, as has often been the case when indicators have been developed until now. A more systematic approach to selecting indicators would assist practitioners when assessing their SDG performance. Moreover, it would enhance harmonization and comparability between indicator assessments made across countries or companies, thus reducing the risk of unintentional bias and misleading communication.

To address this need, the present study proposes a framework for evaluating and selecting SDG indicators in a systematic and transparent manner, which can be adapted to different scales unlike previous efforts that were limited to specific contexts. This is achieved through the following specific objectives: i) mapping the current use of criteria for selecting SDG and sustainability indicators by critically reviewing scientific and gray literature; ii) developing a framework for guiding the evaluation and selection of SDG indicators at different levels; and iii) applying the framework on two proofs-of-concept cases to test its operability. Eventually, the use of the framework should enable the selection and definition of SDG indicators as well as allow for checking existing sets of SDG indicators with regard to their consistency and comprehensiveness.

2. Methods

2.1. Methodological approach

Fig. 1 illustrates the overall methodology used to arrive at the suggested framework and the two main outcomes. In the first step, a literature review was undertaken to identify studies that assessed SDG or sustainability performance and suggested or applied criteria for indicator evaluation. The criteria were retrieved from the identified studies and gathered in a 'pool' of unrefined criteria (see Section 2.2). In the second step, we analyzed the pool of criteria by sorting the list (e.g., aligning the names of similar criteria) and categorizing them to obtain a recommended set of criteria for SDG indicators (i.e., Outcome 1 in Fig. 1, see Section 2.3). The final list consists of mutually exclusive criteria while still comprehensively covering all main aspects proposed in the literature. A framework was developed in the third step based on the findings from the criteria review. It uses Outcome 1 as the main input and proposes a systematized process for evaluating or selecting indicators and reporting hereof (i.e., Output 2 in Fig. 1, see Section 2.4). Finally, to provide input to refine the framework, it was applied to two different illustrative cases: 1) the UN SDG indicators representing a global level assessment and 2) a hypothetical company case representing an organizational level assessment, respectively (see Section 2.5). It is important to note that the proposed framework focuses solely on the indicator selection process and is intended as a support to a more overarching SDG assessment methodology (which is considered outside the scope of this study: see "Full SDG assessment" scoping in Fig. 1).

2.2. Literature review of SDG indicator criteria

A systematic literature review was carried out using Google Scholar (scholar.google.com) and Web of Science (webofscience.com) to identify scientific papers. The review aimed to identify a diverse and representative selection of the literature that mentions criteria for indicator selection. Given the extensive body of literature on this topic, we focused on covering a broad range of studies that applied or discussed criteria across different dimensions of time, scale, and sectors. To scope the search, we defined keywords for identifying studies that either applied or proposed criteria or methods for selection, evaluation, or development of SDG indicators, thus consisting of different combinations of the words 'criteria', 'SDGs', and 'indicators' and relevant synonyms and variants hereof (see more details in Supporting Information 1; i.e., SI-1). Only a few papers were identified targeting SDGs directly; thus, it was decided to expand the search beyond the SDG framework to include criteria for selecting sustainability indicators. The search targeted only publications in English and no temporal restriction to allow for drawing on sustainability assessments that go well past the SDGs. Gray literature was identified using the google search engine, limiting the scope to sources that explicitly referred to the SDGs.

Generally, different synonyms for the word 'criteria' were identified across the studies, hereunder 'requirements', 'principles', 'filters', 'guiding questions', 'characteristics', and 'qualities', which all referred to criteria for evaluating indicators. Additionally, the word criteria was used differently across studies, respectively, as 1) requirements for being suitable/appropriate indicators and 2) overarching sustainability aspects to cover (e.g., economic, environmental, and social). In this paper, we use the word 'criteria' as the former definition, also being the most frequently applied in the identified studies. The term, therefore, covers all the synonyms identified across the studies. In the identification of relevant studies, we only considered the studies that explicitly mentioned proposed or applied criteria. The criteria identified from the studies were gathered into a pool of individual criteria. In this pool, each 'occurrence' of a criterion was documented along with the information of the documenting article, i.e., the same criterion occurring multiple times across studies counts as multiple occurrences. The full list of studies and the pool of criteria can be found in the Supporting

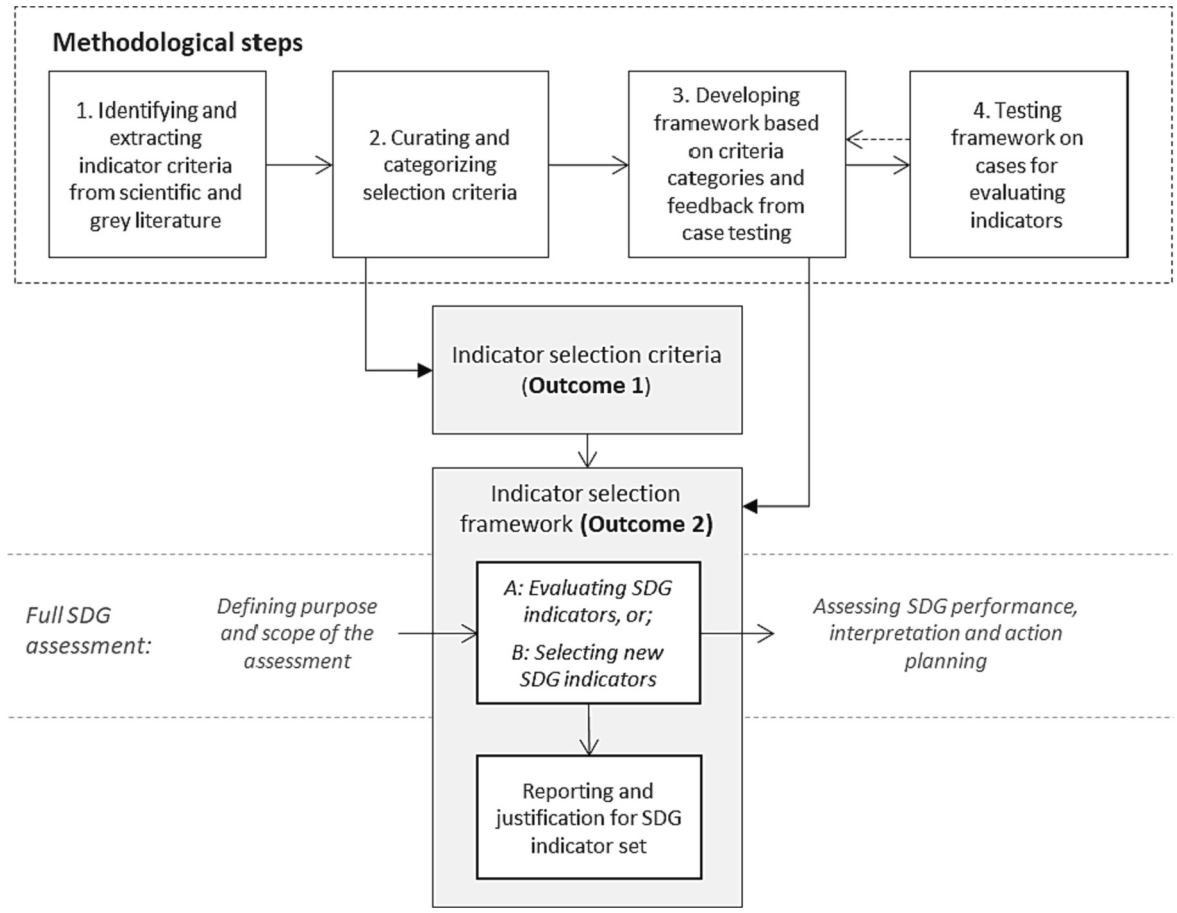

Fig. 1. Methodology steps applied in the development of the indicator selection framework (top) and presentation of the main outcomes; 1) criteria list and 2) indicator selection framework (gray boxes). The vertical gray box (i.e., Outcome 2) defines the scope of the suggested framework of the present study limited to indicator selection. The horizontal steps at the bottom are for contextual purposes only and show the illustrative steps for an SDG-based sustainability performance assessment; they are not considered within the scope of this paper.

Information 2 (i.e., SI-2 Table S1).

2.3. Curation and categorization of criteria list

Different variants of the same criteria were organized by defining them under common names to avoid redundancy in the list (e.g., aligning suffixes: the words relevance, relevancy, and relevant are all the same criterion). Even though a broad set of criteria can ensure a high validity of the indicator set, having too many criteria will increase the complexity and reduce transparency (Pires et al., 2020). Therefore, to arrive at a shortlist, the criteria were categorized based on overarching criteria aspects inspired by the literature (e.g., data-oriented, applicability, relevance, etc.), for which each of them contained 2–4 criteria specifying the categories and ensuring representation of the whole set of criteria. Some criteria were entirely excluded from our list due to vague or ambiguous formulations (see SI-2 Table S2). Finally, we distinguish two lists of criteria: 1) criteria for evaluating individual indicator performance (individual indicator criteria) and 2) criteria for evaluating the performance of the indicator set as a whole (indicator set criteria). The final sets of criteria (i.e., both 1 and 2) were used as the main inputs in the suggested indicator selection method (i.e., Output 1 in Fig. 1).

2.4. Framework development

To arrive at a structured approach for evaluating the performance of or for selecting SDG indicators, a stepwise framework was developed to guide the process (i.e., Output 2 in Fig. 1). The framework was built around the three main steps: 1) pre-steps for identifying potential indicators to evaluate/select, 2) evaluating or selecting indicators, and 3) reporting of the resulting indicator set. Step 2, the main step of the framework, was designed based on findings from the literature review. Here, we distinguished two hierarchal groups of criteria, namely primary/mandatory criteria and secondary/recommended criteria motivated by findings from the review (Mascarenhas et al., 2012, 2015; Tanguay et al., 2013). The former group covers criteria that must be met for indicators to be accepted, and the latter refers to criteria that are less important but can contribute to a more nuanced evaluation of the strength of the indicators. Having mandatory criteria for the indicator selection increases the consistency and scientific soundness while keeping other criteria as recommended allows for a degree of flexibility when applying the criteria.

In line with the objective of this study, the framework is intended to be generically applicable to all SDGs, regardless of the scope of the SDG assessment and whether some SDGs could be deemed irrelevant. These assessment-specific considerations are outside the scope of the paper.

2.5. Illustrative cases

Two cases were chosen to test and validate the applicability of the framework. To assess the applicability at different levels, we demonstrated the framework with i) the proposed UN SDG indicators for a global level assessments and (ii) the selection of SDG indicators for a hypothetical company case (with a focus on selected SDGs in both cases).

Table 1
List of indicator selection criteria for single indicators. Further description of the criteria is available in SI-1. The last column indicates the level of importance of the criteria, where Level A = mandatory for indicators to comply and Level B = recommended for indicators to comply to the furthest extent possible. The listed criteria stem from previous studies that have been retrieved and reviewed through the extensive literature review; detailed references are available in SI-1 Table S1.

Criterion	Description	Evaluation checks[a]	Level
Relevance			
1. Relevant to the scope	The indicator is relevant to the context and the issue being asked (e.g., spatial and temporal scope, area of concern).	Yes: (1) It is relevant to the spatial and temporal scope; and (2) It is relevant to the area of concern/sector/topic (hereunder magnitude and influence on what is measured); and (3) it is relevant to the target audience No: if not yes	**A**
2. Relevant to the SDG	The indicator is relevant to the context of the SDG. Thus, it is important to clearly interpret the SDG in the given context.	Yes: It targets the SDG directly or one or more aspects that constitute the SDG of consideration (e.g., as identified from interpreting the SDG and its targets in the given context) No: if not yes	**A**
General quality			
3. Measurable	The indicator can be measured either quantitatively or semi-quantitatively (e.g., by a categorical or binary scale).	Yes: (1) It can be measured quantitatively; or (2) if not directly, it can be measured by a binary or categorical measure No: if not yes	**A**
4. Performance-based	The indicator measures the performance towards the SDG or a relevant aspect of the SDG. If the indicator measures something further back in the cause-effect chain, e.g., a driver or a more action-oriented indicator or response to the performance level, this does not measure the performance and thus does not comply with the criterion.	Yes: (1) It relates to the performance of the whole or parts of the SDG; and (2) It does not lead or prescribe actions to reach/improve the SDG performance No: if not yes	**A**
5. Scientifically robust	The indicator is based on a scientifically robust method of assessment.	Yes: (1) No cause-effect estimation modeling is required to assess the indicator performance; or (2) If a cause-effect estimation modeling is required, the assessment method should be based on a scientific pathway backed by the scientific community No: if not yes	**A**
6. Comparable	The indicator is comparable across time, space and field. E.g., ensure comparability across companies (for corporate assessments), regions (for global and national assessments), years, etc., given the possibility of normalization.	Yes: (1) It is comparable across time, e.g., not affected by external factors that are linked to the time; and (2) It is comparable across geographical areas (relevant within the scope); and (3) it is comparable across other actors within the scope of the assessment, e.g., countries/ regions/ organizations/ sectors No: if not yes	**A**
Data specific			
7. Data quality	The indicator is measured using high-quality data, i.e., the data is from a reliable, trustworthy, and sound source, and it should be adequately documented.	Yes: (1) The data comes from a reliable/ sound/trustworthy source; and (2) The data is accurate and precise Partly: (1) The data comes from a reliable/ sound/trustworthy source; and (2) the data serves as a proxy for more precise data No: (1) Not applicable since there is no data; or (2) the data is not from a reliable source	**B**
8. Data availability	The data used for measuring the indicators is easily accessible and can be obtained with limited human and financial resources (e.g., calculations, processing).	Yes: (1) The required data is available to the user; and (2) the data does not require any cost or extra calculations and processing Partly: (1) The required data is partly available; and (2) the data requires some effort to achieve or process (e.g., calculations, collection process, etc.) No: (1) The required data is not available; or (2) the data requires substantial effort (e.g., processing/calculations/high costs)	**B**
Acceptability			
9. Broadly accepted	The indicator is generally accepted by involved parties (e.g., stakeholders, end-users). Acceptance should be seen in relation to the scope, e.g., agreement across stakeholders for a company level assessment or agreement across nations in global level assessment.	Yes: (1) It is generally accepted by the involved actors (e.g., stakeholders, users, etc.) at the level of the scope Partly: (1) It is accepted by a large part of the involved actors relevant to the scope No: (1) Only accepted by a few of the involved actors relevant to the scope	**B**
10. Compliant and consensus-based	The indicator should comply with existing systems or standards where these already exist or based on a general consensus within the field of use. New indicators can comply if used for updating standards or developing new standards.	Yes: (1) It complies with existing standards or practices within the field/sustainability area; or (2) There is a general consensus about the use of the indicator within the field of use at a higher level; or (3) if the indicator is new and will be used for updating current standards or suggesting new standards within the field or topic which will be coordinated at a higher level (e.g., sector, country) Partly: (1) It complies with existing standards or practices among some actors within the field/area; or (2) There is some consensus about the use of the indicator across actors	**B**

(*continued on next page*)

Table 1 (*continued*)

Criterion	Description	Evaluation checks[a]	Level
		No: (1) It does not comply with any existing standards or practices within the field; or (2) There is no consensus on the indicator nor a plan to achieve it.	
Applicability			
11. Clear and understandable	The indicator is understandable and unambiguous, uses clear language, and can be understood by users, stakeholders, and policymakers.	Yes: (1) It is easy to understand. I.e., uses language that can be understood by end users, stakeholders, and policymakers; and (2) It is clear how the indicator should be interpreted (e.g., what the desired direction towards sustainability is)	**B**
		Partly: (1) It can be understood by people with some knowledge about the topic; and/or (2) The indicator is hard to interpret	
		No: (1) The indicator is not well defined/described or uses language which is hard to understand by lay people; and (2) The indicator is hard to interpret	
12. Transparent	The indicator is transparently documented and can be replicated or is self-explaining.	Yes: (1) The indicator is self-explaining; or (2) The methodology for assessing it is well documented and easy to replicate	**B**
		Partly: (1) The indicator is not self-explaining; or (2) Documentation for explaining the methodology is provided but needs improvements	
		No: (1) The indicator is not self-explaining; and (2) No documentation for explaining the methodology is provided	

[a] The evaluation checks for the Level A criteria is a binary yes/no check. Since it is mandatory to fulfill these criteria, they require a stricter compliance level. Thus, indicators are only accepted if the check complies with 'yes'. The evaluation checks for the Level B criteria consist of the three options 'yes', 'partly', 'no', allowing for a more flexible evaluation as the indicators are not strictly required to fulfill these criteria.

2.5.1. The UN SDG indicators (global level)

The UN SDG indicators were developed by the Inter-Agency and Expert Group on SDG Indicators (IAEG-SDGs) and agreed upon in 2017 at the 48th session of the United Nations Statistical Commission and in their latest form, in 2022 at the 53rd session (United Nations, 2023; United Nations, 2017). The intention behind the indicators is to monitor the global progress across all SDGs and to assist the work towards achieving the SDGs. The UN SDG indicators are meant to monitor global progress as well as national contributions to inform policymakers at both international and national levels.

Since the UN SDG indicators are already proposed as a comprehensive list for monitoring the SDG progress, the framework will be used solely for evaluation, not a selection of new indicators. Also, only indicator sets for SDG 1 and 13 were evaluated to meet the purpose of the proof-of-concept; these two SDGs are considered of different nature (i.e., social and environmental focus), hence their selection to test the more generic applicability of the framework.

2.5.2. Company indicators with SDG compass (organizational level)

A hypothetical case at the organizational level was defined to illustrate the usability of the framework for selecting new indicators. The case reflects a Danish manufacturing company that provides mechanical and electronic solutions, which is referred to as 'DanMES'. The headquarters are located in Denmark, while the main production facilities are located in Denmark, China, and the US. Smaller facilities and offices are further placed in other European countries.

The company seeks to identify SDG performance indicators to monitor the sustainability performance of the company and to define effective KPIs that can 1) assess the status of the company and 2) motivate improvements. The company has not suggested any SDG indicators previously; thus, the framework should be used for the selection of SDG indicators. The SDG Compass database for business indicators was used as the main basis for the selection process for preparing an indicator test set (GRI et al., 2020). The database was filtered for each of the SDGs, and each indicator was initially screened with regard to the topic and exclusion of duplicates. The assessment was limited to SDG 8, 12, and 13 for a proof-of-concept covering both societal and environmental aspects.

3. Results and discussion

3.1. SDG indicator criteria

A total of 137 studies from the review were deemed relevant by either suggesting or applying indicator criteria. The resulting pool of indicator criteria holds more than 800 occurrences from studies across the period 1988 to 2023 and at various spatial and sectoral scales (see Table S1 in SI-2). Only a few papers from the scientific literature review focused on assessing either a single SDG or SDGs as a whole. However, none of these proposed SDG-specific approaches for the use of criteria when selecting indicators. Furthermore, their use of criteria were based on generic indicator criteria, except from the criteria 'relevance to SDG' and 'relevance to SDG target', thus indicating a lack of systematic approaches for defining SDG indicators (see further details on the review and criteria collection in SI-1). After curating the occurrences and aligning criteria names, the list of criteria was reduced to 80 individual criteria, albeit with overlapping meanings (see the reduced list in SI-2 Table S2). The resulting proposed sets of criteria are presented in Table 1 and Table 2, consisting of 12 criteria for individual indicator performance (described in Section 3.1.1) and 6 criteria for evaluating the set of indicators as a whole (described in Section 3.1.2), respectively (a further description of the indicators can be found in SI-1). Two levels, reflecting the importance of the criteria, are also proposed for the criteria (i.e., Level A and B); these are introduced in Section 3.1.3.

3.1.1. Criteria for individual indicator performance (individual indicator criteria)

Table 1 presents the individual indicator criteria. Here, the first two criteria define the relevance of the indicator for the specific context of the study and for the SDG being considered. Relevance was one of the most frequently mentioned criteria identified from the review and often the first criterion being checked when evaluating indicators (e.g., Bonisoli et al., 2018; Pires et al., 2020). Here, we distinguish between relevance to the spatial or sectoral scope and relevance for the considered SDG.

Another category is the general quality criteria. Like relevance, these cover intrinsic properties of the indicators. First, the indicator should be

Table 2
List of indicator selection criteria for a set of indicators. Further description of the criteria is available in SI-1. The last column indicates the level of importance of the criteria, where Level A = mandatory for indicator sets to comply and Level B = recommended for indicator sets to comply to the furthest extent possible. The listed criteria stem from previous studies that have been retrieved and reviewed through the extensive literature review; detailed references are available in SI-1 Table S1.

Criterion	Description	Evaluation checks[a]	Level
Mutually exclusive			
1. No redundancy	Each indicator in the set provides a unique and non-overlapping function. Each indicator is designed to capture specific facets of the subject matter, and there is no overlap or duplication of information among them.	Yes: Each indicator contributes with its own function and value to the set without bringing repetition of the information within the set (i.e., no overlapping) No: if not yes	A
2. Avoid double-counting in the cause-effect chain	The set of indicators should avoid covering indicators in different places of the cause-effect chain leading to the same effect, which leads to double counting of the same issue.	Yes: The set does not comprise indicators that aim to measure the same effect from different places in the cause-effect chain No: if not yes	A
3. Reasonable size	The number of indicators in the indicator set should be kept at a reasonable number to keep the set focused and manageable.	Yes: The set consists of relatively few indicators per SDG, with a clear focus and low complexity Partly: The set consists of several indicators with some level of complexity without being overwhelming. No: The set consists of a lot of indicators per SDG, which is overwhelming and very complex	B
Collectively exhaustive			
4. Cover the whole life cycle/value chain	The indicator set should consider life cycle thinking when relevant. Together, the set of indicators should link to the whole life cycle/ value chain of the assessed system (e.g., from extraction and manufacturing to use and end of life, or Scope 1, 2, and 3).	Yes: (1) The set consists of indicators that target all scopes of the life cycle; or (2) if it can be well justified that a scope/life cycle stage is not important for the assessment, it may be omitted but this should be explicitly communicated No: if not yes	A
5. Cover of all relevant SDG aspects	The indicator set includes indicators that capture the SDG comprehensively without leaving out certain aspects and potentially overlooking trade-offs (e.g., burden shifting). When trade-offs exist, they should be reflected in the set.	Yes: The set consists of indicators that target all the main aspects of the SDG of consideration with relevance to the given context No: if not yes	A
6. Enable comparison with absolute sustainability targets	The set consists of indicators that can enable to target absolute sustainability at a level that is relevant to the SDGs of consideration. Most indicators can be assessed in terms of relative sustainability (e.g., comparing performance over time and space), however, not all allow for benchmarking against an absolute sustainability reference. For the more outcome-oriented goals, this includes indicators that can compare to consensus-based, external, and independently defined target values. For the more transformative goals; the set includes indicators that target aspects that are proven to leverage the action towards absolute sustainability.	Yes: Where relevant, the set includes indicators for which external, independent, and consensus-based targets or thresholds are defined with which the indicators can be benchmarked against Partly: Where relevant, the set includes indicators for which targets or thresholds are defined with some degree of consensus with which the indicators can be benchmarked against No: The set does not consist of any indicators for which external science-based or policy-based targets or thresholds are defined, i.e., only internally defined targets or thresholds can be defined	B

[a] The evaluation checks for the Level A criteria is a binary yes/no check. Since it is mandatory to fulfill these criteria, they require a stricter compliance level. Thus, indicator sets are only considered if the check complies with 'yes'. The evaluation checks for the Level B criteria consist of the three options 'yes', 'partly', 'no', allowing for a more flexible evaluation as the indicator set is not strictly required to fulfill these criteria.

measurable in order to be monitored (e.g., Jägerbrand, 2021; Karnauskaite et al., 2019; Silva et al., 2020). Often, quantitative indicators are preferred, although in some cases, this might not be possible, or qualitative measures might be more representative (e.g., studies on individual perception or behavior in social sciences). Following the scope of this paper (i.e., SDG performance assessment), the indicator should also be performance-based, meaning indicating the performance of the SDG or parts of it, also referred to as outcome-based in the literature (e.g., Moller and Macleod, 2013; SDSN, 2015). Thus, indicators prescribing actions for how to change performance are not considered suitable. Moreover, the indicator should be scientifically robust (also referred to as credible, valid, reliable, accurate, etc.) (e.g., Asmelash and Kumar, 2019; Boyd and Charles, 2006). In this study, this means that if a cause-effect estimation model is needed for assessing the indicator, it should be based on methods supported by the scientific community. The last criterion in this group concerns comparability (e.g., Casini et al., 2019; Feil et al., 2015; UNCTAD, 2019). To tackle current issues with the lack of comparability between studies, this criterion is proposed to ensure that the indicators are comparable within the context of consideration. For instance, if the assessment of the indicators in a company is being compared to other company performance, the used indicators should be comparable across the companies.

The data-specific criteria were also mentioned frequently in the literature as these are key for the operability of the indicators (e.g., Bonisoli et al., 2018; Pires et al., 2020). Here, we distinguish two criteria: data quality (e.g., Clark et al., 2020; Krank and Wallbaum, 2011) and data availability (e.g., Mvongo et al., 2021; Valizadeh and Hayati, 2021). The former refers to the certainty and trustworthiness of the data and the data source, and the latter refers to whether the data is available to the user. The data-specific criteria are important for the feasibility of using the indicators; nevertheless, it should not come at the expense of selecting strong indicators (as judged by the other criteria). Instead, data quality and availability should be prioritized for improving indicators, while proxy data can be used until better data exists.

The acceptability of the indicators can be useful for strengthening the use of the indicators since it can create a stronger feeling of ownership and involvement by the intended users and a stronger acceptance of the results of the SDG assessment. Therefore, we recommend that indicators should be broadly accepted (e.g., Moller and Macleod, 2013) and compliant with existing systems or standards (e.g., Casini et al., 2019; SDSN, 2015).

Finally, regarding the two applicability criteria, it is recommended that indicators are clear and understandable to the users (e.g., if the indicators are intended for the public, they should not only be

understandable to experts in the field; e.g., Asmelash and Kumar (2019)) and transparently documented (i.e., the indicator used can be reproduced and interpreted by others arriving at the same results; e.g., Reid and Rout (2020); Sobhani et al. (2022)).

Drawing from existing literature, our final proposed list comprehensively covers commonly used criteria identified in the review. Well-established international sources commonly recommend criteria like relevance, specificity, measurability, soundness, and time-boundness for individual indicators (GRI et al., 2020; OECD, 2001; SDSN, 2015). These criteria have strong justifications and are included in our list. However, variations across studies highlight the need for a more comprehensive approach. Notably, these frequently cited reports primarily focus on criteria for individual indicators and often overlook criteria for indicator sets. Criteria for the indicator set are crucial for developing a comprehensive framework to effectively measure SDG performance.

3.1.2. *Criteria for evaluating whole sets of indicators (indicator set criteria)*

In Table 2, the criteria specify indicator sets to follow the principle of being mutually exclusive and collectively exhaustive (i.e., the MECE principle) (González-cabán et al., 1995; Steen and Palander, 2016). As this overarching principle covers several more specific criteria from the literature review, we decided to keep it as two overall categories (i.e., ME and CE) for other criteria. For the mutually exclusive criteria, this category highlights the importance of removing redundancy in the list of indicators (e.g., Mameli and Marletto, 2014; Petrova-Antonova and Ilieva, 2018; Yuan and Lo, 2020). This includes the need for removing indicators providing the same information and indicators targeting the same thing from different places in the cause-effect chain. Having multiple indicators targeting the same information on SDG performance would lead to double-counting and over-representation of the same issue. Finally, it is recommended to keep the number of indicators as low as possible to ease their use and reduce the complexity of the result interpretation (e.g., Krank and Wallbaum, 2011; Moller and Macleod, 2013; Reed et al., 2006).

For the set to be collectively exhaustive, we define three criteria. Firstly, it is important to consider the whole life cycle or value chain of the scope of consideration when measuring sustainability performance (e.g., Feil et al., 2015; Kianian et al., 2018; Kravchenko et al., 2020). Therefore, life cycle coverage is defined as one of the criteria to ensure that indicator sets comprise indicators for all life cycle stages (i.e., including the upstream, operation, and downstream processes of a system). Ideally, one needs to identify how each of the different SDG aspects relates to each of the life cycle stages; however, in practice, this can be a complex task. Therefore, we propose this criterion as a strict minimum condition, where a set consists of at least one indicator in each stage. It should be noted that for some scopes, it is not relevant to consider a life cycle, such as a global assessment of poverty, and, therefore, the set complies with this criterion by default. The second criterion in the category is about comprehensively covering the aspects of the SDG that are relevant to the context, motivated by the criteria for sustainability coverage (e.g., Huang et al., 2015; Silva et al., 2020; Zinkernagel et al., 2018). Here, it is essential to identify the important pieces of the SDG in order to avoid overlooking certain aspects or cherry-picking certain elements (e.g., omitting the environmental dimension of SDG 7 'sustainable energy for all'). The SDG targets can be used to identify these themes and consider their relevance in the given context. For example, SDG 1, Target 1.1, proposes to eradicate extreme poverty. Although this might not be relevant to all contexts, the overall topic 'poverty' could still be targeted across many contexts (e.g., can everyone afford our service?). As for the aforementioned criterion, this criterion is included to force the user to systematically go through the SDG goal definitions and targets and identify indicators that can cover all relevant aspects.

Finally, we recommend considering indicators that enable comparison with absolute sustainability boundaries or targets. Unlike relative sustainability, which compares the performance of different activities (i.e., answering: is it more or less sustainable?), absolute sustainability evaluates whether an activity is inherently sustainable by measuring its performance against predefined absolute sustainability limits (i.e., answering: is it sustainable or not?) (Bjørn et al., 2016; Hauschild, 2015). This criterion was motivated by the result of the review, where some studies suggested indicators for which target or threshold values exist (Moller and Macleod, 2013; Mvongo et al., 2021; Reed et al., 2006). Furthermore, previous studies have pointed out the need for more science-based approaches when assessing the SDGs, hereunder linking to the planetary boundaries and the safe and just space for humanity (Allen et al., 2021; Gebara and Laurent, 2022; Holden et al., 2017). Thus, to accommodate this, we suggest including a criterion that ensures that indicators can be linked to absolute sustainability targets of relevance within each SDG. The methods for defining absolute sustainability targets differ across the sustainability areas, e.g., societal and environmental, where social thresholds can refer to a limit of deprivation based on human needs (e.g., the social foundation; Raworth, 2012), and environmental thresholds can refer to reference values for ecosystem tipping points (e.g., the Planetary Boundaries framework (Steffen et al., 2015)). Thus, different methods should be applied depending on the type of SDG. SDGs, which relate solely to one sustainability dimension (i.e., either social or environmental), should thus be distinguished from SDGs that focus more on societal transformations (e.g., SDG 7 'ensuring access to clean and sustainable energy for all' supporting both environmental and socio-economic aspects; Gebara and Laurent, 2022). In the current study, this criterion simply suggests indicators to enable this comparison; we do not prescribe how such targets should be defined for the given context. It is recommended to consider this step in more detail when developing full SDG assessments. Nevertheless, ensuring indicators that are compatible with variables for measuring absolute sustainability thresholds increases the credibility of the indicator and allows both stakeholders and externals to compare the indicator performance to a relevant absolute reference value at larger scale (i.e., global or national) or define tailored thresholds based on a choice of sharing principle (e.g., allocated share to a sector or company). See further discussion in Section 3.4 and further description in SI-1.

3.1.3. *Criteria hierarchy*

Inspired by the observations from the identified studies, the framework proposes two hierarchal groupings of criteria (cf. Section 2.4), namely mandatory (i.e., strict compliance) and recommended criteria (i.e., compliance to the furthest extent possible), referred to respectively as 'Level A' and 'Level B' criteria (see last column in Table 1 and Table 2). In Table 1, the criteria were divided into the two levels based on the nature of the criteria, namely intrinsic (i.e., referring to the inherent qualities of the criterion) and extrinsic (i.e., referring to the operability of the indicator). The relevancy and general quality criteria are intrinsic criteria. If an indicator does not fulfill these criteria, it is irrelevant to consider the remaining criteria since it already disqualifies. Therefore, the indicator is only accepted if it answers 'yes' to the evaluation checks in Table 1. The remaining criteria on data, acceptance, and applicability are more extrinsic in nature, and they can, therefore, be improved and do not necessarily disqualify the eligibility of the indicator (e.g., as more data becomes available, the indicator can improve). Furthermore, as there can be situations where some criteria can be hard to satisfy in parallel with other criteria (e.g., ensuring scientific robustness, which might increase the complexity level and limit data availability), it is instead recommended to fulfill these extrinsic criteria to the largest extent possible. For these criteria, the evaluation checks, therefore, consists of a scale, where the indicators can be evaluated as 'yes', 'partly', and 'no', representing their degree of compliance.

For the indicator-set criteria, all the criteria were kept as mandatory (i.e., Level A), except for 'indicator of reasonable size' and 'comparability with absolute sustainability targets'. The reason for keeping most criteria as Level A was to ensure a strong set of SDG indicators as a whole, not just at individual indicator level. Thus, to adhere to more

scientifically based indicator sets, these criteria are deemed strictly necessary. Yet, we acknowledge that adhering to coverage of all life cycle stages and all SDG aspects are not straightforward criteria and can be complex to address. Therefore, we apply them by setting minimum requirements for "good practice" for complying as a first step towards addressing these needs. Nevertheless, the operability of these criteria can be tackled differently in future studies. The two criteria of Level B were kept at this level in this study due to 1) difficulties in setting an exact number of indicators a set should contain, as this depends on the scope, and 2) the lack of maturity in the criterion for linking to absolute sustainability thresholds and defining strict requirements of compliance.

3.2. Proposed framework for SDG indicators

Two variants of the framework are suggested depending on the purpose of its application: Variant 1, when used for indicator evaluation (i.e., after the definition and selection of indicators), and Variant 2 for indicator selection (i.e., prior to defining a new set of indicators) (see Outcome 2 in Fig. 1). The two variants follow the same overall structure and build on the same sets of criteria. They are, however, presented separately on two figures to underline their different uses, where Variant 1 (Fig. 2) is a more simplified version that allows for quick screening of the indicator quality of an existing set, and Variant 2 (Fig. 3) is a more extended version of the former and requires some extra steps and active choices by the user when developing a new SDG indicator set. Both variants consist of a pre-step and five following main steps. The pre-steps (i.e., in the dotted box) cover the needs for framing the project and preparing a set of indicators prior to applying the framework. This includes defining the main goal and scope of the entity to be assessed. Another pre-step is to prepare the pool of indicators to test. For Variant 1, indicators are already defined, and the user can go directly to the application of the framework. For Variant 2, it is necessary to prepare a list or a set of sources as the basis for selecting new indicators. Since this is very much linked to the scope of the project (e.g., national, sectoral, or company-level assessment), the recommended sourcing of indicators can vary a lot, and different sources might exist for different levels of assessments (e.g., the inventory of business indicators by the SDG Compass for the level of company assessments; GRI et al., 2020). Thus, identifying different sources for indicators within the field of consideration is an important step, hereunder, both internally and externally sourced indicators.

The framework iterates over each of the SDGs, starting with SDG 1 (i.e., X = 1), followed by SDG 2 (i.e., X = 1 + 1), and so on. Similarly, it iterates over each indicator, starting with Indicator X.1 (i.e., Y = 1),

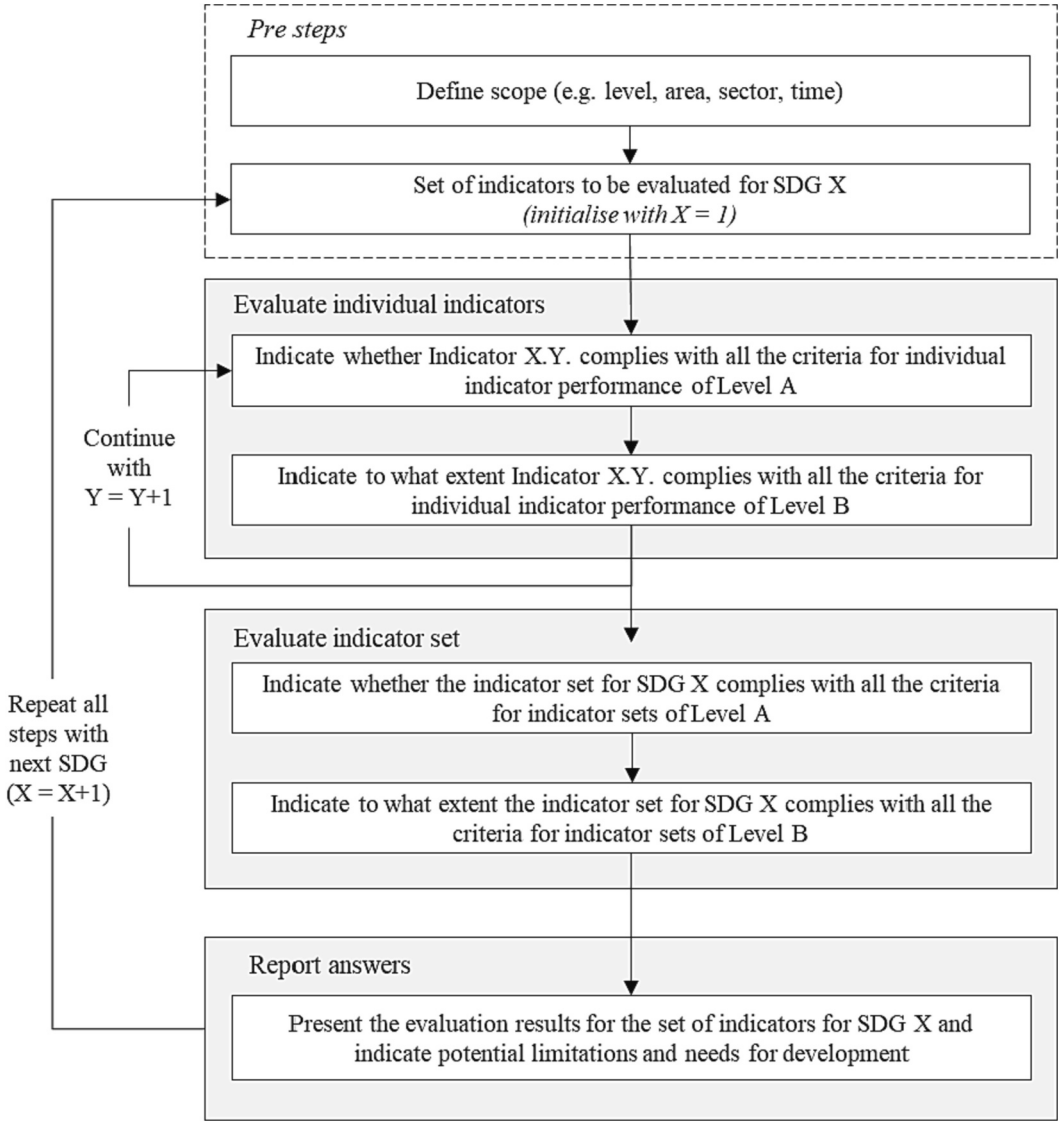

Fig. 2. Framework for SDG indicator evaluation (Variant 1).

Indicator X.2 (i.e., Y = 1 + 1), and so on. Each indicator for an SDG is evaluated against the Level A criteria to qualify or disqualify individual indicators, while Level B criteria are used to identify potential limitations and motivate further improvement. As an example of framework application to an individual indicator, we take the indicator "amount of GHG emissions" for measuring SDG 13 performance of the agricultural sector. It is considered relevant to the context and the SDG (clear link to SDG-13), is measurable (quantifiable), focuses on performance (lower/higher values indicate progress or not), it can be assessed with a scientifically backed method (e.g., carbon footprinting), and it is comparable (either by total magnitude or in relative terms when normalized). Thus, the indicator meets all Level A criteria, and the Level B criteria can thereafter be evaluated. These criteria are more context-specific; in a hypothetical example, assuming the stakeholders have access to data and have had an including process for selecting the indicator, it is assumed that the indicator aligns with all criteria, except data quality, due to lack of information about upstream emissions and downstream handling. To apply our framework correctly, the users should document the justification for the alignment with each criterion and report that the criterion on data availability is only partly met, highlighting the limitations and needs for improvement. Since all Level A criteria are met and the Level B criteria have been checked and reported, the indicator could be kept as a potential candidate to keep in the final set. The criteria for the indicator set is then used to evaluate which indicators to keep in the final set once a potential set of candidates has been identified, again prioritizing the Level A criteria followed by the evaluation against the Level B criteria. Finally, the framework includes a step of reporting the choices and justifications of how the indicators comply with all the criteria, hereunder, the potential development needs and limitations of the indicator set. Since choosing indicators will always have some

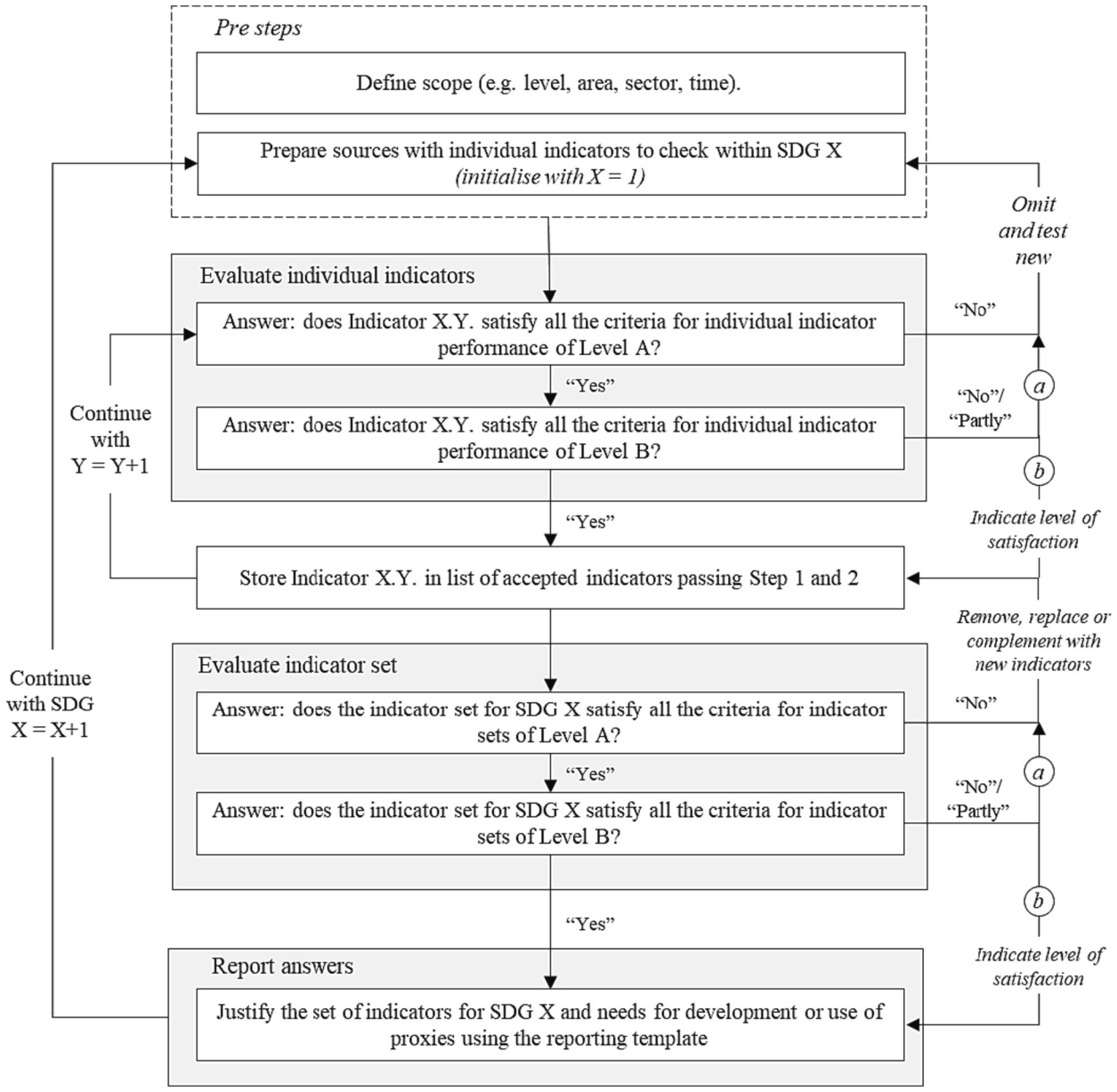

Fig. 3. Framework for SDG indicator selection (Variant 2). The small circles with 'a' and 'b' indicate options for handling indicators and indicator sets when they do not meet or partly meet the requirements of the Level B criteria. The 'a' circle indicates an option to omit (when assessing individual indicators), remove, replace, or complement with new indicators (when assessing a set of indicators) if indicators or a set of indicators do not meet the requirements of the criteria. The 'b' circles indicate the option to continue if they do not or partly meet the requirements of the criteria. X indicates the SDG number, and Y indicates the indicator number for the given X.

degree of subjectivity, documenting and communicating transparently about the choices can reduce the bias of the results.

The framework does not ensure one final solution of indicators, as the choice of indicators will always have some degree of subjectivity. Furthermore, the set of indicators is highly dependent on the scope of the assessment and the starting point for the indicator evaluation or selection. Nevertheless, it sets the general frame with a set of minimum requirements for indicators to follow across all scopes and enables a guide for disqualifying poor indicators and indicator sets. The framework is intended as a key input for improving the indicator selection process in a full SDG performance assessment and can be a starting point for developing more sector-specific methods.

3.3. Framework implementation on cases

3.3.1. Case 1: global and national assessment

Indicators were evaluated for SDG 1 and SDG 13 using Variant 1 of the framework. The evaluation is presented in Fig. 4 as a heat map, indicating the degree of compliance: 'yes', 'partly', or 'no', by the colors green, yellow, and red, respectively. A more detailed evaluation for each indicator and criterion can be found in SI-2 Table S3 and S4.

Overall, the indicators of both SDGs fulfill most criteria. All indicators were deemed relevant to both the scope and the SDGs. However, for the general indicator quality category, some indicators conflict with the criteria of being measurable, performance-based, and comparable. For example, the indicator 1.2.2. "Proportion of men, women and children of all ages living in poverty in all its dimensions according to national definitions" was not considered compliant with the criterion of being measurable due to the non-tangible formulation of "all dimensions". For this criterion to be fulfilled, this term should be further specified in the indicator definition to avoid ambiguity and misinterpretation.

The three indicators 1.5.3/13.1.2. "Number of countries that adopt and implement national disaster risk reduction strategies in line with the Sendai Framework for Disaster Risk Reduction 2015–2030", 1.5.4./13.1.3. "Proportion of local governments that adopt and implement local disaster risk reduction strategies in line with national disaster risk reduction strategies", and 13.2.1 "Number of countries with nationally determined contributions, long-term strategies, national adaptation plans and adaptation communications, as reported to the secretariat of the United Nations Framework Convention on Climate Change" are not performance-based as they all focus on the implementation of policy strategies and thus prescribe certain actions to improve on the performance. Therefore, they were evaluated as not fulfilling Criterion 4 in the list of individual indicator criteria (i.e., Table 1). The judgment of dismissing these indicators from the SDG performance indicator set does not suggest that they are irrelevant. Instead of being treated similarly to the rest of the indicators, these indicators should be seen as complementary potential action-based indicators for improving performance, not as performance indicators in themselves. Thus, we suggest treating

Fig. 4. Indicator evaluation of SDG 1 and SDG 13 for Case 1 'UN SDG indicators' using the indicator framework Variant 1. The level of compliance with the criteria is presented by the colors 'green', 'yellow', and 'red' indicating 'yes', 'partly', 'no' respectively. The single rows on the right side represent the indicator evaluation for the whole set. More elaborated results can be found in SI-2 Table S3 and S4.

these as the means of implementation indicators of the UN SDG indicators (not assessed in the scope of this paper).

The following indicators do not fulfill the criterion for comparability: 1.2.1. "Proportion of population living below the national poverty line, by sex and age", 1.2.2. "Proportion of men, women and children of all ages living in poverty in all its dimensions according to national definitions", and 13.3.1 "Extent to which (i) global citizenship education and (ii) education for sustainable development are mainstreamed in (a) national education policies; (b) curricula; (c) teacher education; and (d) student assessment" as they all refer to national context or reference values, which disqualifies their comparability across countries (cf. SI-2 Table S3 and S4).

All indicators performed well with respect to compliance with existing systems and broad acceptance (i.e., Criterion 9 and 10 in Table 1). The main hotspots were identified with respect to data availability criterion and to a lesser extent, clarity and transparency. Low data availability is a common limitation of the UN SDG indicators, as data in many countries is either not generated or estimated too seldom (Schmidt-Traub et al., 2017). The issues with clarity and transparency are linked to indicators where concepts can be interpreted differently by different users (see SI-2 for detailed evaluation).

With respect to indicator set criteria, the main issues appear for the criteria of being mutually exclusive, hereunder avoiding redundancy and overlapping in the cause-effect chain. An example is Indicator 1.1.1. "Proportion of population below the international poverty line (…)" and 1.2.1. "Proportion of population living below the national poverty line", which are both measuring poverty level. Another is Indicator 1.5.3. "Number of countries that adopt and implement national disaster risk reduction strategies" and 1.5.4. "Proportion of local governments that adopt and implement local disaster risk reduction strategies" both measuring disaster risk strategies. Moreover, measuring economic losses due to disaster can be seen as a cause of increasing the poverty level; thus, both are tackling poverty eradication from different places of the cause-effect chain.

Even though the indicator evaluation (i.e., Variant 1 of the framework) does not imply developing a new UN SDG indicator set in the current assessment, the result allows for an overview of the main limitations in existing indicators/indicator sets. Thereby, it proposes where to target improvements and prioritize indicators to develop or replace (i.e., red-marked indicators for Level A in Fig. 4).

3.3.2. *Case 2: organizational assessment*

Applying Variant 2 of the framework for this case resulted in a proposed set of indicators for SDG 8, 12, and 13, presented in Table 3. The framework was applied based on the indicator sourcing using the SDG Compass database (see method in Section 2.5.2).

The indicators of the final sets all pass the Level A criteria (i.e., mandatory), which is why these criteria are merged into the same column in Table 3. Other relevant indicators were initially identified from the SDG Compass database but were removed as they were focusing on policies rather than the actual performance of an SDG aspect (i.e., not performance-based). However, for the Level B (i.e., recommended) criteria, some indicators were not fully compliant or only partly compliant (i.e., indicated by yellow or red in Table 3). These were mainly linked to data availability for the company and transparent documentation of the indicator application (i.e., Criterion 7, 8, and 12 in the list of individual indicator criteria) due to the assumed difficulty with gathering data for Scope 3 activities (i.e., upstream and downstream activities excl. energy purchases). Its determination is, therefore, complex and is often the source of limitations for organizations (Stenzel and Waichman, 2023). For the set of SDG 13, some indicators were only judged as partly fulfilling the criterion of compliance and consensus as different methodologies for calculating Scope 3 emissions for companies are used across assessments (Schmidt et al., 2022). It should be noted that the answers here are fully hypothetical, nevertheless showcasing how the criteria can be applied for such a case.

Making sure to arrive at a suite of indicators that covered well the life cycle, the indicators were classified with regards to which life cycle stage or scope of the company they were targeting, i.e., upstream (Scope 3), operations (typically mainly Scopes 1 and 2), and downstream (Scope 3) (see SI-2 Table S5-S7 for more details). This enabled identifying if all scopes were addressed and whether some were underrepresented. Moreover, it was noted which aspects of the SDG the indicators were targeting (e.g., diversity, working conditions, or economy for SDG 1). These steps assisted the selection process of the final indicator sets to make sure that all life cycle stages and SDG aspects were considered (see SI-2 Table S5-S7). For the criterion of linking to absolute sustainability, the sets of SDG 8 and 12 were evaluated as partly compliant (i.e., yellow in Table 3) because they include indicators that can relate to absolute sustainability targets for a single sustainability dimension (e.g., zero child labor and forced labor). However, a deeper analysis of the nature of these more transformative SDGs (i.e., not solely targeting environmental or social end goals) should be carried out as more indicators could be relevant, tapping into multiple sustainability dimensions. For SDG 13, the indicator sets were deemed compliant as the SDG focuses on the environmental dimension (i.e., climate change) as an outcome and includes an indicator – namely the total GHG emissions – that can be related to absolute biophysical limits (Friedlingstein et al., 2022; Hickel, 2020). Finally, all three sets were deemed reasonable in size, therefore making it easier to assess and interpret the SDG performance by the whole suite.

Some indicators were selected for both SDG 12 and SDG 13, i.e., identical indicators for both SDGs to capture direct and indirect GHG emissions (i.e., Indicators 21–23 in Table 3). These overlaps are a part of the interconnected nature of the 17 SDGs, as also seen in the UN SDG indicator framework, where identical indicators are used as metrics for different SDGs (e.g., the indicator on 'material footprint' is used for both SDG 8 and SDG 12). However, to avoid double counting and overweighting of the same indicators, we assigned Indicators 21–23 to SDG 13 and present their link to SDG 12 in parenthesis in Table 3. If SDG 12 is assessed alone, the three indicators should, nonetheless, be included to ensure coverage of all aspects of the SDG. Other approaches could be considered to account for these overlaps as long as the interlinkages are transparent and double-counting is avoided. The proposed set for SDG 12 aims to cover all aspects of the SDG except "sustainable practices", which mainly refer to more action-based indicators prescribing a certain practice, and it was therefore not considered suitable for a performance assessment. As for the case of SDG 12, some aspects of SDG 13 were also left out of the set concerning education and awareness. These were not included since they focused on the means of mitigating climate change and not on the actual performance. Nevertheless, relevant indicators for awareness raising and education are still recommended to complement the set of performance-based indicators to identify means of action and whether the actions are leading to the desired effect.

Based on the framework steps, an indicator set was successfully selected, which can be used to assess the current status of DanMES (i.e., answering the needs of the company). The indicators fulfill all the Level A criteria. For the Level B criteria, some indicators can be improved to reach even better compliance. To achieve this, DanMES should initiate the task of sourcing more data and present and elaborate on how the indicators should be measured and interpreted in the context of their assessment. This will help the company to improve the indicators with regard to the Level B criteria. Furthermore, the company should focus on improving data and information about suppliers and identify potential action points to improve performance. However, this lies out of the scope of this assessment.

3.3.3. *Usefulness of the framework*

The indicator set in Table 3 showcases the operability of the framework for the selection of SDG indicators at company level. The indicators should, however, not be seen as generic for other companies, nor should they be considered the sole 'solution' to deriving SDG

Table 3
Proposed indicator set for SDG 8, 12, and 13 for Case 2 and the evaluation of the indicator qualities. The numbers in the columns to the right refer to each of the criteria from Table 1 and Table 2. The criteria of Level A are merged into one column for each of the two lists since all indicators comply with these, being a prerequisite for selecting them. **The color coding represents: green = yes, yellow = partly, and red = no.** A full description of the indicators, source of documentation, and criteria assessment are detailed in SI-2 Table S5-S7.

Proposed indicators for Case 2	SDG	Theme	Individual indicator criteria							Set criteria		
			1-6	7	8	9	10	11	12	1,2 4,5	3	6
1. Ratio of basic salary and remuneration of women to men	8	Diversity										
2. Percentage of employees per employee category in each of the following diversity categories: i. Gender; ii. Age group: iii. Other indicators of diversity where relevant	8	Diversity										
3. Total number of employees that took parental leave, by gender	8	Diversity, rights										
4. Operations and suppliers considered to have significant risk for incidents of i. Child labor; ii. Young workers exposed to hazardous work	8	Working conditions										
5. Operations and suppliers considered to have significant risk for incidents of forced labor	8	Working conditions										
6. Total number of incidents of discrimination at company facilities	8	Working conditions										
7. Injury rate and work-related fatalities by type, for all workers whose work, or workplace, is controlled by the organization, with a breakdown by:i. Region;ii. Gender	8	Working conditions										
8. Workers with high incidence or high risk of diseases related to their occupation	8	Working conditions										
9. Investments in activities with risk of child labor, forced labor, violated workers' rights (yes/no)	8	Working conditions										
10. Direct economic value generated and distributed on an accrual basis.	8	Economy										
11. Whether the company's buying practices impact price volatility of key commodities and/or inputs that suppliers rely on in local or national markets (yes/no)	8	Economy										
12. Resource use per economic value generated (by type of resource)	8	Resource intensity										
13. GHG emissions per economic value generated	8	Resource intensity										
14. Amount of significant air emissions, incl. NO_x, SO_x, VOCs, hazardous air pollutants, PM, ODSs, etc.	12	Emissions										
15. Energy consumption by energy sources for all scopes possible	12	Resource use										
16. Energy intensity at different life cycle stages or scopes	12	Resource use										
17. Water consumption: i) total consumption (by scope), ii) intensity (by output), iii) share of recycling and reuse, iv) amount discharged	12	Resource use										
18. Trends in Ecological Footprint and/or related concepts (material or land footprint)	12	Resource use										
19. Materials used breakdown by weight or volume	12	Resource use										
20. Total amount of wastes (incl. water discharge) by type and disposal method	12	Wastes										
21. Direct greenhouse gas (GHG) emissions (Scope 1) in CO2 equivalent.	13(12)	Emissions										
22. Energy indirect greenhouse gas (GHG) emissions (Scope 2) in CO2 equivalent	13(12)	Emissions										
23. Other indirect greenhouse gas (GHG) emissions (Scope 3) in CO2 equivalent	13(12)	Emissions										
24. Avoided emissions from application of product/service (either in CO2 equivalent or binary yes/no)	13	Emissions										
25. Emission intensities - company's gross global combined Scope 1 and 2 (and Scope 3 if possible) emissions in metric tonnes CO2e per unit currency total revenue	13	Emission intensity										
26. Emission intensities - company's gross global combined Scope 1 and 2 (and Scope 3 if possible) emissions for the reporting year in metric tonnes CO2e per full-time equivalent (FTE) employee	13	Emission intensity										
27. Investments in fossil fuel activities or practices that contribute to increased GHG emissions (yes/no)	13	Emissions										

indicators for a company like DanMES. Furthermore, it is likely that two practitioners will not arrive at the exact same set of indicators from applying the framework, as the starting point and justification of the practitioner's choices are decisive factors. Nevertheless, it offers a systematic approach for evaluating a set of indicators against a consistent and comprehensive system of quality requirements (i.e., the criteria) and supports improvement and selection of indicator sets that are grounded in a scientific framework. Consequently, it caters to a higher degree of comparability between indicator sets when applied by different persons, unlike previous work, where indicators are selected without explicit criteria or where criteria are used arbitrarily.

The two proof-of-concept cases provided above illustrate how a screening of indicators can be carried out systematically and as a way to quickly spot limitations of the indicator sets and identify focus points for improvement. It should be kept in mind that the company case used in this study is a fully hypothetical case. Therefore, the case does not build on actual data availability nor a real selection process among stakeholders. Nonetheless, it serves as a representative model applicable to real-world scenarios of similar scope. The selected indicators have been designed with generality in mind, making them suitable for a wide range of manufacturing companies. This adaptability is a result of the criteria for comparability and performance orientation, avoiding overly specific indicators tailored to a single company. Consequently, the set of indicators we have developed represents a robust framework for assessing SDG performance across diverse companies. With that said, when applying the framework to real cases, the indicator may be evaluated very differently with regards to the Level B criteria dependent on the company (e.g., due to actual data availability or stakeholder involvement). It should additionally be noted that the application of the criteria to the virtual proof-of-concept bears limitations in the feedback it can offer, and it is recommended to build experience in applying the framework on real-life case studies, which would provide additional valuable information, such as the user-friendly nature or the full relevance of the results for indicator selection or definition. Such feedback could help refine the list of criteria and facilitate its integration into a wider methodology for SDG performance assessment.

Although only a few SDGs were tested here, the framework suggests generic applicability, and thus, the same approach can be reproduced for the remaining SDGs (i.e., those not included in this study). Reporting openly and transparently about the choices and justifications behind the selection of indicators reduces the risk of misinterpretations when results are communicated. Thus, while there might be different arguments for judging an indicator as being relevant to the context (i.e., Criterion 1 in Table 1), this should be communicated so that the reader understands the reasoning behind it and thus uncovers otherwise hidden bias. Furthermore, it allows stakeholders and other reviewers to identify poorly justified or arbitrary choices when they are not well-backed or can be criticized. However, communicating openly about the limitations of certain selected indicators should not be used as complacency but as a focus point for improving the assessment.

Key limitations are associated with our framework. Firstly, it does not guarantee uniform indicator selection among different individuals. Previous work has argued for this inconsistency mainly due to the use of different indicator sets. Different approaches for reducing this risk have been proposed, such as having a unified and comprehensive indicator base to support the selection process (Warchold et al., 2022), classifying indicator systems by their nature and purpose (Lepenies et al., 2023), or with the use of indicator factsheets to support policymakers in making

informed selections (Hák et al., 2015). Our framework would benefit strongly by being complemented by such approaches in the selection process to address subjectivity. To address this limitation further, we emphasize the importance of transparent documentation, which exposes the subjectivity embedded in selecting indicator sets and serves as a means of motivation. Nevertheless, to mitigate this issue further, it may be valuable to develop more specific guidelines or evaluation checks tailored to distinct entities or scopes, such as detailed requirements to define relevance in a certain sector (see Section 3.5).

Additionally, some of our criteria are not as fully developed or operationalized as others, such as the coverage of SDG aspects and the linkage with absolute sustainability targets. To enhance the framework's robustness, it would be beneficial to engage in further discussions among relevant research scholars to refine and solidify these criteria.

3.4. Relevance for criteria on absolute sustainability perspective

To overcome current issues of arbitrary target-setting and subjective choices of SDG indicators, the framework proposes the need to include an absolute sustainability perspective. This is done via the inclusion of Criterion 6 in the indicator set criteria (i.e., Table 2). When addressing absolute sustainability, two aspects are important to consider, namely 1) what is being measured and 2) which target or threshold value is being used. The criterion proposed in this study aims to address the first aspect by ensuring that the included indicators are relevant with regard to being benchmarked against absolute reference values. Setting the target values relies highly on the choice of tolerance level (e.g., how much environmental impact can we allow?). At smaller scale (e.g., company level), it depends on ethical sharing principles (e.g., who should have what share of the safe operating space within the tolerance level?), for which consensus does currently not exist (Ryberg et al., 2020). However, prior to defining these targets, it is important to start defining what we need to measure if the performance should allow for benchmarking against absolute sustainability references.

Absolute sustainability should not be seen as a replacement for relative sustainability but rather as complementary. Having both allows for measuring improvement over time and how far we are from becoming sustainable in absolute terms. The criterion suggested in this study does not strictly recommend the extent to which the sets of indicators can relate to absolute sustainability, as the concept is still in its developing phase, and what defines absolute sustainability across the SDGs has different levels of complexity (e.g., what is the absolute sustainability of cities?). Further research and consensus are needed to answer such questions for all 17 SDGs.

3.5. Relevance for context-specific criteria

Answering previous calls for science-based and informed support for SDG assessment (Allen et al., 2021; Hák et al., 2016), this framework attempts to define an approach that ensures a degree of objectivity, comparability, and scientific foundation between performance assessments. Thus, the possibility of selecting very activity-specific indicators is limited in the currently proposed framework, which keeps the indicator assessment focused on the SDG performance outcomes. Nevertheless, it can be useful to consider more specific indicators as complementary to the SDG indicator sets, which can support or suggest context-specific actions towards improving the SDG indicators. The reason for not including this option in the criteria within this study is that such specific indicators should not be considered as SDG indicators but rather as indicators for guiding towards actions that can improve SDG performance. E.g., if a company wants to improve the outcome-based indicator ‘amount of GHG emissions’, the hypothetical indicators ‘weight of materials used in products’ or ‘expenses on imported products’ might be relevant for identifying action points for reducing GHG emissions. Here, it is important to consider internally actionable indicators within their value chain rather than those outside of it (van Zanten and van Tulder, 2021). Even so, they should not represent the SDG performance per se as they do not indicate how well we are doing with regard to the SDG. Furthermore, we risk overlooking other aspects contributing to the SDG.

While the proposed framework suggests a basis and starting point for selecting an indicator set, extending the framework further could be to propose fixed sets of indicators that apply across similar activities (e.g., sectoral or regional standards). Here, it can be an advantage to develop ‘core’ and ‘optional’ indicators (i.e., fulfilling the criteria of different types) when designing an indicator set within a certain sector or project type (Karnauskaite et al., 2019). This can improve the comparability of studies even further and, at the same time, allow for tailor-made indicator solutions where the optional indicators can be chosen more freely.

4. Conclusions and Recommendations

A comprehensive list of 18 criteria for SDG indicator selection was proposed in the current study, consisting of two sets: one for individual indicator performance and the other for indicator set performance, and a hierarchy of two levels, i.e., Level A and Level B for mandatory and recommended criteria, respectively. The proposed framework demonstrated its operability in the two cases respectively for Variant 1 and Variant 2, and provides a first step towards a systematic framework for applying those criteria for SDG indicator selection (both retrospective and in the development phase). The framework successfully enabled the assessment of both evaluating existing global level indicators (i.e., the UN SDG indicators) and guided the selection process for company level indicators answering previous research calls. In order to achieve a balance between objective and scientifically rigorous indicators while still accommodating flexibility and tailor-made indicators, the framework advocates for a mix of a top-down and bottom-up approach. This approach defines objective criteria while allowing the practitioner the freedom to select the indicators to be tested. Unlike previous works, the specificity of our study is that we compile and integrate a list of criteria into a structured approach that specifically caters to the SDG framework. Taking existing lists would inherently run the risk of overlooking relevant criteria or including criteria that may not be relevant to the SDG framework because too context-specific.

Some limitations of the current framework were discussed. To tackle these, we suggest the following research needs for improving the framework and its implementation:

i. improving the guidance and definition of the criteria that need further development (i.e., linking to absolute sustainability and other system-oriented criteria)
ii. testing on real-life cases at different levels to collect experience from different applications of the framework, like its ease-of-use and its actual added value in indicator selection/definition, to help refine and support development of a full-fledged methodology for SDG performance assessment.
iii. developing context-specific frameworks/guidelines for indicator selection at different scales or sectors
iv. developing robust frameworks/guidelines for performing full-sized SDG performance assessments (e.g., including well-defined scope definition, data collection, target-setting, and interpretation of indicator performance) that integrate the current framework for indicator selection step
v. developing context-specific mandatory indicators or standards to improve comparability across studies within the same scale/context of consideration

Despite these needs being directly linked to the development of SDG assessments, some broader research needs remain to be solved. Hereunder, understanding the SDG interlinkages, the need for weighting of certain SDG aspects, and the development of methods for defining absolute sustainability targets at different levels. In this paper, we

proposed a framework that acts as a first step towards a general guide that can be adopted in the existing SDG assessment framework or used for developing new tools. The framework should not stand alone but should be integrated into a full-fledged SDG assessment method. As the need for an SDG assessment might differ a lot across the scope of the study, such frameworks could be tailored to the specific scale of concern. The indicator selection remains a complex and decisive step in sustainability assessment. Thus, presenting this first guide can lay the foundation for further improvements and tailored context-specific assessments.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.spc.2023.12.004.

References

Allen, C., Metternicht, G., Wiedmann, T., 2018. Initial progress in implementing the Sustainable Development Goals (SDGs): a review of evidence from countries. Sustain. Sci. 13, 1453–1467. https://doi.org/10.1007/s11625-018-0572-3.

Allen, C., Reid, M., Thwaites, J., Glover, R., Kestin, T., 2020. Assessing national progress and priorities for the Sustainable Development Goals (SDGs): experience from Australia. Sustain. Sci. 15, 521–538. https://doi.org/10.1007/s11625-019-00711-x.

Allen, C., Metternicht, G., Wiedmann, T., 2021. Priorities for science to support national implementation of the sustainable development goals: a review of progress and gaps. Sustain. Dev. 29, 635–652. https://doi.org/10.1002/sd.2164.

Asmelash, A.G., Kumar, S., 2019. Assessing progress of tourism sustainability: developing and validating sustainability indicators. Tour. Manag. 71, 67–83. https://doi.org/10.1016/j.tourman.2018.09.020.

Bjørn, A., Margni, M., Roy, P.O., Bulle, C., Hauschild, M.Z., 2016. A proposal to measure absolute environmental sustainability in life cycle assessment. Ecol. Indic. 63, 1–13. https://doi.org/10.1016/j.ecolind.2015.11.046.

Bonisoli, L., Galdeano-Gómez, E., Piedra-Muñoz, L., 2018. Deconstructing criteria and assessment tools to build agri-sustainability indicators and support farmers' decision-making process. J. Clean. Prod. 182, 1080–1094. https://doi.org/10.1016/j.jclepro.2018.02.055.

Boyd, H., Charles, A., 2006. Creating community-based indicators to monitor sustainability of local fisheries. Ocean Coast. Manag. 49, 237–258. https://doi.org/10.1016/j.ocecoaman.2006.03.006.

Cagno, E., Neri, A., Howard, M., Brenna, G., Trianni, A., 2019. Industrial sustainability performance measurement systems: a novel framework. J. Clean. Prod. 230, 1354–1375. https://doi.org/10.1016/j.jclepro.2019.05.021.

Casini, M., Bastianoni, S., Gagliardi, F., Gigliotti, M., Riccaboni, A., Betti, G., 2019. Sustainable development goals indicators: a methodological proposal for a multidimensional fuzzy index in the mediterranean area. Sustain 11, 1198. https://doi.org/10.3390/su11041198.

Clark, C.M.A., Kavanagh, C., Lenihan, N., 2020. Measuring Progress: The Sustainable Progress Index 2020. Soc. Justice Irel.

Feil, A.A., de Quevedo, D.M., Schreiber, D., 2015. Selection and identification of the indicators for quickly measuring sustainability in micro and small furniture industries. Sustain. Prod. Consum. 3, 34–44. https://doi.org/10.1016/j.spc.2015.08.006.

Field, C.B., Barros, V., Stocker, T.F., Dahe, Q., Jon Dokken, D., Ebi, K.L., Mastrandrea, M. D., Mach, K.J., Plattner, G.K., Allen, S.K., Tignor, M., Midgley, P.M., 2012. Managing the risks of extreme events and disasters to advance climate change adaptation: special report of the intergovernmental panel on climate change. In: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation: Special Report of the Intergovernmental Panel on Climate Change. https://doi.org/10.1017/CBO9781139177245.

Friedlingstein, P., O'sullivan, M., Jones, M.W., Andrew, R.M., Gregor, L., Hauck, J., Le Quéré, C., Luijkx, I.T., Olsen, A., Peters, G.P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J.G., Ciais, P., Jackson, R.B., Alin, S.R., Alkama, R., Arneth, A., Arora, V.K., Bates, N.R., Becker, M., Bellouin, N., Bittig, H.C., Bopp, L., Chevallier, F., Chini, L.P., Cronin, M., Evans, W., Falk, S., Feely, R.A., Gasser, T., Gehlen, M., Gkritzalis, T., Gloege, L., Grassi, G., Gruber, N., Gürses, Ö., Harris, I., Hefner, M., Houghton, R.A., Hurtt, G.C., Iida, Y., Ilyina, T., Jain, A.K., Jersild, A., Kadono, K., Kato, E., Kennedy, D., Klein Goldewijk, K., Knauer, J., Korsbakken, J.I., Landschützer, P., Lefèvre, N., Lindsay, K., Liu, J., Liu, Z., Marland, G., Mayot, N., Mcgrath, M.J., Metzl, N., Monacci, N.M., Munro, D.R., Nakaoka, S.I., Niwa, Y., O'brien, K., Ono, T., Palmer, P.I., Pan, N., Pierrot, D., Pocock, K., Poulter, B., Resplandy, L., Robertson, E., Rödenbeck, C., Rodriguez, C., Rosan, T.M., Schwinger, J., Séférian, R., Shutler, J.D., Skjelvan, I., Steinhoff, T., Sun, Q., Sutton, A.J., Sweeney, C., Takao, S., Tanhua, T., Tans, P.P., Tian, X., Tian, H., Tilbrook, B., Tsujino, H., Tubiello, F., Van Der Werf, G.R., Walker, A.P., Wanninkhof, R., Whitehead, C., Willstrand Wranne, A., Wright, R., Yuan, W., Yue, C., Yue, X., Zaehle, S., Zeng, J., Zheng, B., 2022. Global Carbon Budget 2022. Earth Syst. Sci. Data 14, 4811–4900. https://doi.org/10.5194/essd-14-4811-2022.

Gebara, C.H., Laurent, A., 2022. National SDG-7 performance assessment to support achieving sustainable energy for all within planetary limits. Renew. Sust. Energ. Rev. 173, 112934 https://doi.org/10.1016/J.RSER.2022.112934.

González-cabán, A., Fenn, M.E., Scatena, F.N., 1995. Concepts, Criteria, and Indicators for Monitoring Sustainability 13–23.

GRI, UNGC, WBCSD, 2020. Inventory of Business Tools – SDG Compass [WWW Document]. URL. https://sdgcompass.org/business-indicators/ (accessed 3.15.23).

Hák, T., Janoušková, S., Whitby, A., Abdallah, S., Kovanda, J., 2015. Indicator policy factsheets: a knowledge brokerage tool. Sustain 7, 3414–3429. https://doi.org/10.3390/SU7033414.

Hák, T., Janoušková, S., Moldan, B., 2016. Sustainable development goals: a need for relevant indicators. Ecol. Indic. 60, 565–573. https://doi.org/10.1016/j.ecolind.2015.08.003.

Hauschild, M.Z., 2015. Better - but is it good enough? On the need to consider both eco-efficiency and eco-effectiveness to gauge industrial sustainability. In: Procedia CIRP. Elsevier B.V. 29, 1–7. https://doi.org/10.1016/j.procir.2015.02.126.

Heras-Saizarbitoria, I., Urbieta, L., Boiral, O., 2022. Organizations' engagement with sustainable development goals: from cherry-picking to SDG-washing? Corp. Soc. Responsib. Environ. Manag. 29, 316–328. https://doi.org/10.1002/csr.2202.

Hickel, J., 2020. Quantifying national responsibility for climate breakdown: an equality-based attribution approach for carbon dioxide emissions in excess of the planetary boundary. Lancet Planet. Heal. 4, e399–e404. https://doi.org/10.1016/S2542-5196(20)30196-0.

Holden, E., Linnerud, K., Banister, D., 2017. The imperatives of sustainable development. Sustain. Dev. 25, 213–226. https://doi.org/10.1002/sd.1647.

Huang, L., Wu, J., Yan, L., 2015. Defining and measuring urban sustainability: a review of indicators. Landsc. Ecol. 30, 1175–1193. https://doi.org/10.1007/s10980-015-0208-2.

IPCC, 2022a. Summary for Policymakers. Climate Change 2022 Mitigation of Climate Change Working Group III Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.

IPCC, 2022b. Climate Change 2022: Mitigation of Climate Change, Working Group III Contribution to the Sixth Assessment Report (AR6), Intergovernmental Panel on Climate Change (IPCC).

Jägerbrand, A.K., 2021. Development of an indicator system for local governments to plan and evaluate sustainable outdoor lighting. Sustain 13, 1–22. https://doi.org/10.3390/su13031506.

Janoušková, S., Hák, T., Moldan, B., 2018. Global SDGs assessments: helping or confusing indicators? Sustain 10 (1540 10), 1540. https://doi.org/10.3390/SU10051540.

Karnauskaite, D., Schernewski, G., Støttrup, J.G., Kataržyte, M., 2019. Indicator-based sustainability assessment tool to support coastal and marine management. Sustain 11. https://doi.org/10.3390/su11113175.

Kianian, B., Daly, E., Andersson, C., 2018. Towards guidelines for selection of production performance indicators to measure sustainability performance. Proc. Manuf. 25, 570–577. https://doi.org/10.1016/j.promfg.2018.06.099.

Kim, K.H., Kabir, E., Ara Jahan, S., 2014. A review of the consequences of global climate change on human health. J. Environ. Sci. Heal. C Environ. Carcinog. Ecotoxicol. Rev. 32, 299–318. https://doi.org/10.1080/10590501.2014.941279.

Krank, S., Wallbaum, H., 2011. Lessons from seven sustainability indicator programs in developing countries of Asia. Ecol. Indic. 11, 1385–1395. https://doi.org/10.1016/j.ecolind.2011.02.017.

Kravchenko, M., Pigosso, D.C.A., McAloone, T.C., 2020. A procedure to support systematic selection of leading indicators for sustainability performance measurement of circular economy initiatives. Sustain 12, 951. https://doi.org/10.3390/su12030951.

Lepenies, R., Büttner, L., Bärlund, I., Jax, K., Lyytimäki, J., Pedersen, A.B., Nielsen, H.Ø., Mosoni, C., Mille, R., Payen, G., Richard, D., 2023. The politics of national SDG indicator systems: a comparison of four European countries. Ambio 52, 743–756. https://doi.org/10.1007/s13280-022-01809-w.

Lu, J., Liang, M., Zhang, C., Rong, D., Guan, H., Mazeikaite, K., Streimikis, J., 2021. Assessment of corporate social responsibility by addressing sustainable development goals. Corp. Soc. Responsib. Environ. Manag. 28, 686–703. https://doi.org/10.1002/csr.2081.

Mameli, F., Marletto, G., 2014. Can national survey data be used to select a core set of sustainability indicators for monitoring urban mobility policies? Int. J. Sustain. Transp. 8, 336–359. https://doi.org/10.1080/15568318.2012.700000.

Mascarenhas, A., Ramos, T.B., Nunes, L., 2012. Developing an integrated approach for the strategic monitoring of regional spatial plans. Land Use Policy 29, 641–651. https://doi.org/10.1016/j.landusepol.2011.10.006.

Mascarenhas, A., Nunes, L.M., Ramos, T.B., 2015. Selection of sustainability indicators for planning: combining stakeholders' participation and data reduction techniques. J. Clean. Prod. 92, 295–307. https://doi.org/10.1016/j.jclepro.2015.01.005.

Mengistu, A.T., Panizzolo, R., 2023. Tailoring sustainability indicators to small and medium enterprises for measuring industrial sustainability performance. Meas. Bus. Excell. 27, 54–70. https://doi.org/10.1108/MBE-10-2021-0126.

Millward-Hopkins, J., Steinberger, J.K., Rao, N.D., Oswald, Y., 2020. Providing decent living with minimum energy: a global scenario. Glob. Environ. Chang. 65, 102168 https://doi.org/10.1016/j.gloenvcha.2020.102168.

Moller, H., Macleod, C.J., 2013. Design criteria for effective assessment of sustainability in New Zealand's production landscapes. NZ Sustain. Dashboard Res. Rep. 13 (07), 73.
Mvongo, V.D., Defo, C., Tchoffo, M., 2021. Sustainability of rural water services in rural sub-Saharan Africa environments: developing a water service sustainability index. Sustain. Water Resour. Manag. 7, 1–17. https://doi.org/10.1007/s40899-021-00526-8.
Niemeijer, D., de Groot, R.S., 2008. A conceptual framework for selecting environmental indicator sets. Ecol. Indic. 8, 14–25. https://doi.org/10.1016/j.ecolind.2006.11.012.
OECD, 2001. OECD Environmental Indicators. OECD Environ. Indic, 25. https://doi.org/10.1787/9789264193499-en.
Petrova-Antonova, D., Ilieva, S., 2018. Smart cities evaluation - a survey of performance and sustainability indicators. In: Proceedings - 44th Euromicro Conference on Software Engineering and Advanced Applications, SEAA 2018. Institute of Electrical and Electronics Engineers Inc., pp. 486–493. https://doi.org/10.1109/SEAA.2018.00084
Pires, A., Morato, J., Peixoto, H., Bradley, S., Muller, A., 2020. Synthesizing and standardizing criteria for the evaluation of sustainability indicators in the water sector. Environ. Dev. Sustain. 22, 6671–6689. https://doi.org/10.1007/s10668-019-00508-z.
Pizzi, S., Rosati, F., Venturelli, A., 2021. The determinants of business contribution to the 2030 Agenda: introducing the SDG reporting score. Bus. Strateg. Environ. 30, 404–421. https://doi.org/10.1002/bse.2628.
Raworth, K., 2012. A safe and just space for humanity: can we live within the doughnut?. In: Oxfam Discussion Paper. State World 2003 Prog. Towar. a Sustain. Soc. 20th Ed, pp. 1–240.
Reed, M.S., Fraser, E.D.G., Dougill, A.J., 2006. An adaptive learning process for developing and applying sustainability indicators with local communities. Ecol. Econ. 59, 406–418. https://doi.org/10.1016/j.ecolecon.2005.11.008.
Reid, J., Rout, M., 2020. Developing sustainability indicators – the need for radical transparency. Ecol. Indic. 110, 105941 https://doi.org/10.1016/j.ecolind.2019.105941.
Ryberg, M.W., Andersen, M.M., Owsianiak, M., Hauschild, M.Z., 2020. Downscaling the planetary boundaries in absolute environmental sustainability assessments – a review. J. Clean. Prod. 276 https://doi.org/10.1016/j.jclepro.2020.123287.
Sachs, J.D., Lafortune, G., Fuller, G., Drumm, E., 2023. Implementing the SDG Stimulus. Sustainable Development Report 2023, Sustainable Development Report 2023. Paris: SDSN. Dublin University Press, Dublin (10.25546/102924).
Schmidt, M., Nill, M., Scholz, J., 2022. Determining the scope 3 emissions of companies. Chem. Eng. Technol. 45, 1218–1230. https://doi.org/10.1002/ceat.202200181.
Schmidt-Traub, G., Kroll, C., Teksoz, K., Durand-Delacre, D., Sachs, J.D., 2017. National baselines for the sustainable development goals assessed in the SDG index and dashboards. Nat. Geosci. https://doi.org/10.1038/NGEO2985.
SDSN, 2015. Indicators and a Monitoring Framework for the Sustainable Development Goals Launching a Data Revolution for the SDGs A Report to the Secretary-General of the United Nations by the Leadership Council of the Sustainable Development Solutions Network.
Silva, J. da, Fernandes, V., Limont, M., Rauen, W.B., 2020. Sustainable development assessment from a capitals perspective: analytical structure and indicator selection criteria. J. Environ. Manag. 260, 110147 https://doi.org/10.1016/j.jenvman.2020.110147.
Sobhani, P., Esmaeilzadeh, H., Sadeghi, S.M.M., Marcu, M.V., Wolf, I.D., 2022. Evaluating ecotourism sustainability indicators for protected areas in Tehran, Iran. Forests 13, 740. https://doi.org/10.3390/f13050740.
Steen, B., Palander, S., 2016. A selection of safeguard subjects and state indicators for sustainability assessments. Int. J. Life Cycle Assess. 21, 861–874. https://doi.org/10.1007/s11367-016-1052-6.
Steffen, W., Richardson, K., Rockström, J., Cornell, S.E., Fetzer, I., Bennett, E.M., Biggs, R., Carpenter, S.R., De Vries, W., De Wit, C.A., Folke, C., Gerten, D., Heinke, J., Mace, G.M., Persson, L.M., Ramanathan, V., Reyers, B., Sörlin, S., 2015. Planetary boundaries: guiding human development on a changing planet. Science 347. https://doi.org/10.1126/science.1259855.
Stenzel, A., Waichman, I., 2023. Supply-chain data sharing for scope 3 emissions. npj Clim. Action 2. https://doi.org/10.1038/s44168-023-00032-x.
Sullivan, D., Hickel, J., 2023. Capitalism and extreme poverty: a global analysis of real wages, human height, and mortality since the long 16th century. World Dev. 161, 106026 https://doi.org/10.1016/j.worlddev.2022.106026.
Tanguay, G.A., Rajaonson, J., Therrien, M.C., 2013. Sustainable tourism indicators: selection criteria for policy implementation and scientific recognition. J. Sustain. Tour. 21, 862–879. https://doi.org/10.1080/09669582.2012.742531.
UNCTAD, 2019. Guidance on Core Indicators for Entity Reporting on Contribution Towards Implementation of the Sustainable Development Goals. https://doi.org/10.18356/1902575e-en.
UNFCCC, 2015. Adoption of the Paris Agreement, Conference of the Parties on its Twenty-first Session (doi:FCCC/CP/2015/L.9).
United Nations, 2015. Transforming Our World: The 2030 Agenda for Sustainable Development. https://doi.org/10.1891/9780826190123.ap02.
United Nations, 2019. Annex: Global Indicator Framework for the Sustainable Development Goals and Targets of the 2030 Agenda for Sustainable Development. In: Work Stat. Comm. Pertain. to 2030 Agenda Sustain. Dev, 1–21.
United Nations, 2022. Nations Adopt Four Goals, 23 Targets for 2030 in Landmark UN Biodiversity Agreement. Press Release.
United Nations, 2023. Global indicator Framework for the Sustainable Development Goals and Targets of the 2030 Agenda for Sustainable Development, IAEG-SDGs.
United Nations Statistical Commission, 2017. Report of the Inter-Agency and Expert Group on Sustainable Development Goal Indicators (E/CN.3/2017/2), Annex IV.
Valizadeh, N., Hayati, D., 2021. Development and validation of an index to measure agricultural sustainability. J. Clean. Prod. 280, 123797 https://doi.org/10.1016/j.jclepro.2020.123797.
Warchold, A., Pradhan, P., Thapa, P., Putra, M.P.I.F., Kropp, J.P., 2022. Building a unified sustainable development goal database: why does sustainable development goal data selection matter? Sustain. Dev. 30, 1278–1293. https://doi.org/10.1002/sd.2316.
Yuan, M.H., Lo, S.L., 2020. Developing indicators for the monitoring of the sustainability of food, energy, and water. Renew. Sust. Energ. Rev. 119, 109565 https://doi.org/10.1016/j.rser.2019.109565.
van Zanten, J.A., van Tulder, R., 2021. Improving companies' impacts on sustainable development: a nexus approach to the SDGS. Bus. Strateg. Environ. 30, 3703–3720. https://doi.org/10.1002/bse.2835.
Zinkernagel, R., Evans, J., Neij, L., 2018. Applying the SDGs to cities: business as usual or a new dawn? Sustain 10, 3201. https://doi.org/10.3390/su10093201.